

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 27, 2005, 15:23:02 ; Search time 8691 Seconds  
(without alignments)

11503.997 Million cell updates/sec

Title: US-09-930-591-1

Perfect score: 2061

Sequence: 1 atggcgccatcacggccta.....atgaaatgaagagtgtga 2061

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.\*

1: gb\_ba.\*

2: gb\_hgt.\*

3: gb\_in.\*

4: gb\_om.\*

5: gb\_ov.\*

6: gb\_pat.\*

7: gb\_ph.\*

8: gb\_pl.\*

9: gb\_pr.\*

10: gb\_ro.\*

11: gb\_sts.\*

12: gb\_sy.\*

13: gb\_un.\*

14: gb\_vi.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2061	100.0	2061	6	C0826998 Sequence
2	2061	100.0	2061	6	AX441176 Sequence
3	2061	100.0	2061	6	AX467113 Sequence
4	1854.8	90.0	9610	14	HEC278830
5	1846.4	89.6	8791	14	AY615798
6	1787	86.7	6299	6	AX164584
7	1786	86.7	5360	6	I06434
8	1786	86.7	5360	6	I09328
9	1786	86.7	6785	6	I06440
10	1786	86.7	7310	6	AR118696
11	1786	86.7	7310	6	I09331
12	1786	86.7	7310	14	HPCPOLYP
13	1786	86.7	9185	6	I08294
14	1786	86.7	9185	6	BD091382
15	1786	86.7	9379	6	AR166930
16	1786	86.7	9379	6	AR301300
17	1786	86.7	9401	6	AR176483
18	1786	86.7	9401	6	E66593
19	1786	86.7	9401	6	I71894

20	1786	86.7	9401	6	I81885	I81885 Sequence 9
21	1786	86.7	9401	6	BD080334	BD080334 Hepatitis
22	1786	86.7	9401	14	HPCELYPRE	M62321 Hepatitis C
23	1784.4	86.6	6785	6	AR118692	I09329 Sequence 10
24	1784.4	86.6	6785	6	I09329	I09329 Sequence 10
25	1784.4	86.6	8316	6	AR118703	AR118703 Sequence
26	1784.4	86.6	8987	6	AR118728	AR118728 Sequence
27	1784.4	86.6	9185	6	AR118722	AR118722 Sequence
28	1784.4	86.6	9185	6	AR118723	AR118723 Sequence
29	1784.4	86.6	9379	6	AR118747	AR118747 Sequence
30	1781.2	86.4	2058	6	AR404933	AR404933 Sequence
31	1781.2	86.4	2058	6	AX395309	AX395309 Sequence
32	1781.2	86.4	5360	6	AR118686	AR118686 Sequence
33	1781.2	86.4	9424	14	AF511948	AF511948 Hepatitis
34	1781.2	86.4	9609	12	AF387805	AF387805 Synthetic
35	1781.2	86.4	9618	14	AF271632	AF271632 Hepatitis
36	1781.2	86.4	9646	12	AF387806	AF387806 Synthetic
37	1781.2	86.4	9693	12	AF387807	AF387807 Synthetic
38	1779.6	86.3	2058	6	AR408362	AR408362 Sequence
39	1779.6	86.3	2058	6	AX454818	AX454818 Sequence
40	1774.8	86.1	12980	6	AR110831	AR110831 Sequence
41	1774.8	86.1	12980	6	BD069985	BD069985 Functiona
42	1773.2	86.0	9609	12	AF387808	AF387808 Synthetic
43	1773.2	86.0	9646	6	AR110828	AR110828 Sequence
44	1773.2	86.0	9646	6	BD069982	BD069982 Functiona
45	1773.2	86.0	9646	14	AF009606	AF009606 Hepatitis

ALIGNMENTS

RESULT 1						
LOCUS	C0826998	2061 bp	DNA	linear	PAT 29-JUN-2004	
DEFINITION	Sequence 1 from Patent WO2004048402.					
ACCESSION	C0826998					
VERSION	C0826998.1	GI:49455655				
KEYWORDS	synthetic construct					
SOURCE	synthetic construct					
ORGANISM	other sequences; artificial sequences.					
REFERENCE	1					
AUTHORS	Sallberg,M.					
TITLE	A hepatitis c virus codon optimized non-structural ns3/4a fusion gene					
JOURNAL	Patent: WO 2004048402-A 1 10-JUN-2004;					
FEATURES	TRIPEP AB (SE)					
source	Location/Qualifiers					
	1..2061					
	/organism="synthetic construct"					
	/mol_type="unassigned DNA"					
	/db_xref="taxon:32630"					
	/note="Hepatitis C virus NS3/4A coding region"					

ORIGIN						
Query Match	100.0%;	Score 2061;	DB 6;	Length 2061;		
Best Local Similarity	100.0%;	Pred. No. 0;				
Matches 2061;	Conservative	0;	Mismatches	0;	Indels	0; Gaps 0;
Qy	1	ATGGCGCCTATCAGCGCCTATGCCAGCAGACAAAGGGCCCTTTGGGATGCATATCACC	60			
Db	1	ATGGCGCCTATCAGCGCCTATGCCAGCAGACAAAGGGCCCTTTGGGATGCATATCACC	60			
Qy	61	AGCTTACCGCGCGGACAAAACCAAGGTGAGGTTCAGATCGTGCACTGCT	120			
Db	61	AGCTTACCGCGCGGACAAAACCAAGGTGAGGTTCAGATCGTGCACTGCT	120			
Qy	121	GCCCAGACATTTCTTTGGCACTGCTTACCGGCTGTGTGGGACTGTCTACCATGGAGCC	180			
Db	121	GCCCAGACATTTCTTTGGCACTGCTTACCGGCTGTGTGGGACTGTCTACCATGGAGCC	180			
Qy	181	GGAAACAGGACCAATTCGTCACCTAAGGTCCTGTTATCCAGATGTACCAATGTGGAC	240			
Db	181	GGAAACAGGACCAATTCGTCACCTAAGGTCCTGTTATCCAGATGTACCAATGTGGAC	240			

```
Db 181 GGAACAAGGACCAATTGGCTGACCTAAAGGTCCTGTATCCAGATGTACCAATGTGGAC 240
Qy 241 CAAGACCTCGTAGGCTGGCCGCTCCCAAGGTGCGGCTCATTTAAACCATGACATTGC 300
Db 241 CAAGACCTCGTAGGCTGGCCGCTCCCAAGGTGCGGCTCATTTAAACCATGACATTGC 300
Qy 301 GGCTCCTCGGACCTTTACCTGGTCAAGGACGACCGCATGTCATTCCTGTGCGCGGACGG 360
Db 301 GGCTCCTCGGACCTTTACCTGGTCAAGGACGACCGCATGTCATTCCTGTGCGCGGACGG 360
Qy 361 GGTGATGCGAGGCGAGCTGCTTTTCGCCCCGGGCTATCTCTTAAAGGCTCCCTCG 420
Db 361 GGTGATGCGAGGCGAGCTGCTTTTCGCCCCGGGCTATCTCTTAAAGGCTCCCTCG 420
Qy 421 GGAGGCTCTGCTGTCGCGCGAGGACATGCGGTAGGCATATTCAGAGCGCGGTATGC 480
Db 421 GGAGGCTCTGCTGTCGCGCGAGGACATGCGGTAGGCATATTCAGAGCGCGGTATGC 480
Qy 481 ACCCGTGGAGTGGCTAAGGGGCTGGACTTCATCCCGGTAGAGAGCTTAGAGACAACCATG 540
Db 481 ACCCGTGGAGTGGCTAAGGGGCTGGACTTCATCCCGGTAGAGAGCTTAGAGACAACCATG 540
Qy 541 AGTCCCGGCTGTTCTCAGACAACCTCTCCCAACGACGAGTGCCTCAGAGCTACCAAGTG 600
Db 541 AGTCCCGGCTGTTCTCAGACAACCTCTCCCAACGACGAGTGCCTCAGAGCTACCAAGTG 600
Qy 601 GCCACCTGCATGCTCCCAAGGCGGTGAAGGACCAAGGTCCTCCGCGGATAGCGCA 660
Db 601 GCCACCTGCATGCTCCCAAGGCGGTGAAGGACCAAGGTCCTCCGCGGATAGCGCA 660
Qy 661 GCTCAGGCTACAGGTGCTGCTCAACCCCTCCGTTGCTGCAACAATGGGCTTTGGT 720
Db 661 GCTCAGGCTACAGGTGCTGCTCAACCCCTCCGTTGCTGCAACAATGGGCTTTGGT 720
Qy 721 GCTTACATGTCGAAGGCCATGGGATTGATCTTAAACATCAGGACTGGGGTGAGGACAAAT 780
Db 721 GCTTACATGTCGAAGGCCATGGGATTGATCTTAAACATCAGGACTGGGGTGAGGACAAAT 780
Qy 781 ACTACTGGACCGCGATCAGTATTCACCTACGCAAGTTCCTTCCGACGCGCGGCTGT 840
Db 781 ACTACTGGACCGCGATCAGTATTCACCTACGCAAGTTCCTTCCGACGCGCGGCTGT 840
Qy 841 TCAGGGGCTCTTATGACATAAATTTGTGACGAGTGCCATCCACGATGCAACATCC 900
Db 841 TCAGGGGCTCTTATGACATAAATTTGTGACGAGTGCCATCCACGATGCAACATCC 900
Qy 901 ATCTTGGGCAATGGCACTGCTTACCAAGAGAGACCGCGGGGCGAGACTGACTGTG 960
Db 901 ATCTTGGGCAATGGCACTGCTTACCAAGAGAGACCGCGGGGCGAGACTGACTGTG 960
Qy 961 CTGCGCACCGCTACCCCTCGGGCTCCGTCACTGTCGCCCATCTAACATCGAGGAGTT 1020
Db 961 CTGCGCACCGCTACCCCTCGGGCTCCGTCACTGTCGCCCATCTAACATCGAGGAGTT 1020
Qy 1021 GCTCTGTCCACTACCGAGAGATCCCTTTTATGGCAAGGCTATTCCTCTTGAAGCAAT 1080
Db 1021 GCTCTGTCCACTACCGAGAGATCCCTTTTATGGCAAGGCTATTCCTCTTGAAGCAAT 1080
Qy 1081 AAGGGGGGAGACATCTCATCTTCTGCACTCAAGAGAGTGCAGAGCTCGCGGCA 1140
Db 1081 AAGGGGGGAGACATCTCATCTTCTGCACTCAAGAGAGTGCAGAGCTCGCGGCA 1140
Qy 1141 AAATGCTGCGGTGGGCGTCAATCGCTGGCTTACTACCGGCGCTTGTATGTCCGTC 1200
Db 1141 AAATGCTGCGGTGGGCGTCAATCGCTGGCTTACTACCGGCGCTTGTATGTCCGTC 1200
Qy 1201 ATCCGACAGTGGTGAAGTGTGCTGGGCACTGAGCCCTCATGCCGCTTTTACC 1260
Db 1201 ATCCGACAGTGGTGAAGTGTGCTGGGCACTGAGCCCTCATGCCGCTTTTACC 1260
Qy 1261 GGGGACTTCGATTCGGGTAGACTGCAACAGTGTGTACCAGACAGTCACTTCAG 1320
Db 1261 GGGGACTTCGATTCGGGTAGACTGCAACAGTGTGTACCAGACAGTCACTTCAG 1320
```

```
Qy 1321 CTTGACCCCTACCTTACCAATTGAGACAATCA CGCTTCCCAGGATGCTGTCTCCCGTACT 1380
Db 1321 CTTGACCCCTACCTTACCAATTGAGACAATCA CGCTTCCCAGGATGCTGTCTCCCGTACT 1380
Qy 1381 CAACGTCGGGTAGGACTGCGAGAGGAAGCAGGATCTACAGATTTGTGGCACCGGG 1440
Db 1381 CAACGTCGGGTAGGACTGCGAGAGGAAGCAGGATCTACAGATTTGTGGCACCGGG 1440
Qy 1441 GAGCGCTCTTCTGGCAATGTTTGA CTGCTGCTCTCTGCGAGTGTCTATGACGCGGTTGT 1500
Db 1441 GAGCGCTCTTCTGGCAATGTTTGA CTGCTGCTCTCTGCGAGTGTCTATGACGCGGTTGT 1500
Qy 1501 GCTTGTATGAGCTTAGGCGCGGAGACCA CAGTTAGGCTACGAGCATACATGAACACC 1560
Db 1501 GCTTGTATGAGCTTAGGCGCGGAGACCA CAGTTAGGCTACGAGCATACATGAACACC 1560
Qy 1561 CCGGGACTTCCCGTGTGCCAAGACCATCTTGAATTTTGGAGGGCTCTTTACGGGTCTC 1620
Db 1561 CCGGGACTTCCCGTGTGCCAAGACCATCTTGAATTTTGGAGGGCTCTTTACGGGTCTC 1620
Qy 1621 ACCACATAGACGCCCATTTCTCTATCC CAGACAAAGCAGAGTGGGAAAACCTTCCCTAT 1680
Db 1621 ACCACATAGACGCCCATTTCTCTATCC CAGACAAAGCAGAGTGGGAAAACCTTCCCTAT 1680
Qy 1681 CTGGTAGCGTACCAAGCCACCGTGTGCG CTAGAGCTCAAGCCCTCCCGCTCGTGGAC 1740
Db 1681 CTGGTAGCGTACCAAGCCACCGTGTGCG CTAGAGCTCAAGCCCTCCCGCTCGTGGAC 1740
Qy 1741 CAGATGTGAAGTGTCTGATCCGTCTCAAG CCCCACCTCCATGGGCAACACTCTGCTA 1800
Db 1741 CAGATGTGAAGTGTCTGATCCGTCTCAAG CCCCACCTCCATGGGCAACACTCTGCTA 1800
Qy 1801 TATAGACTGGGCGCTGTCAGAAATGAAGT CACCTGAGCACCAGTACCAGTATATC 1860
Db 1801 TATAGACTGGGCGCTGTCAGAAATGAAGT CACCTGAGCACCAGTACCAGTATATC 1860
Qy 1861 ATGACATGATCTCGGCTGACCTGAGGTGCT CAGAGTACCTGCTGCTTGGCGC 1920
Db 1861 ATGACATGATCTCGGCTGACCTGAGGTGCT CAGAGTACCTGCTGCTTGGCGC 1920
Qy 1921 GTTCTGGCTGCTTTGGCGCGCTATTGCC TATCCAGAGTGGCTGTATAGTAGTAGG 1980
Db 1921 GTTCTGGCTGCTTTGGCGCGCTATTGCC TATCCAGAGTGGCTGTATAGTAGTAGG 1980
Qy 1981 ATTGTCTTGTCCGAAAAGCGGCAATCATC CCGACAGGAAAGTCTCTTACCGGAGTTC 2040
Db 1981 ATTGTCTTGTCCGAAAAGCGGCAATCATC CCGACAGGAAAGTCTCTTACCGGAGTTC 2040
Qy 2041 GATGAAATGGAAGAGTGTCTGA 2061
Db 2041 GATGAAATGGAAGAGTGTCTGA 2061
```

## RESULT 2

```
AX441176
LOCUS AX441176 2061 bp DNA linear PAT 28-JUN-2002
DEFINITION Sequence 16 from Patent WO0213855.
ACCESSION AX441176
VERSION AX441176.1 GI:21665758
KEYWORDS
SOURCE synthetic construct
ORGANISM other sequences; artificial sequences.
REFERENCE
1
AUTHORS Sallberg, M. and Hultgren, C.
TITLE Vaccines containing ribavirin and methods of use thereof
JOURNAL Patent: WO 0213855-A 16 21-FEB-2002;
TRIPEP AB (SE)
FEATURES
source location/Qualifiers
1..2061
/mol_type="synthetic construct"
/mol_type="unassigned DNA"
```

ORIGIN		/db_xref="taxon:32630"		/note="Hepatitis C virus NS3/4A coding region"	
Query Match		100.0%; Score 2061; DB 6; Length 2061;			
Best Local Similarity		100.0%; Pred. No. 0;			
Matches 2061; Conservative 0; Mismatches 0; Indels 0; Gaps 0;					
Qy	1	ATGCGCCTATCAGCGCTATGCCAGCAGACAAGGGGCTTTTGGGATGCATAATCAC	60		
Db	1	ATGCGCCTATCAGCGCTATGCCAGCAGACAAGGGGCTTTTGGGATGCATAATCAC	60		
Qy	61	AGCTTGACCGGCGGGGAAAAACACAGGTGAGGGTGAAGTTCAGATCGTGCATCGCT	120		
Db	61	AGCTTGACCGGCGGGGAAAAACACAGGTGAGGGTGAAGTTCAGATCGTGCATCGCT	120		
Qy	121	GCCAGACCTTTCTTGGCAACTGATTAACGGGTGTTGGAGTGTCTACCATGGAGCC	180		
Db	121	GCCAGACCTTTCTTGGCAACTGATTAACGGGTGTTGGAGTGTCTACCATGGAGCC	180		
Qy	181	GGAAACAGGACCATTTGCGTCACCTTAAGGGTCTGTATCCAGATGTACACCAATGTG	240		
Db	181	GGAAACAGGACCATTTGCGTCACCTTAAGGGTCTGTATCCAGATGTACACCAATGTG	240		
Qy	241	CAAGACCTGTAGCTGGCGCGCTCCCAAGGTGCCCGCTCATTAACACCATGCACTTGC	300		
Db	241	CAAGACCTGTAGCTGGCGCGCTCCCAAGGTGCCCGCTCATTAACACCATGCACTTGC	300		
Qy	301	GGCTCTCGGACCTTTACTGTCACGAGCAGCGCATGCTATCTCTGTCGCGCAGCG	360		
Db	301	GGCTCTCGGACCTTTACTGTCACGAGCAGCGCATGCTATCTCTGTCGCGCAGCG	360		
Qy	361	GGTGATGGCAGGGGACCTGCTTTTCGCCCGGCTATCTTACTTTGAAAGGCTCCTCG	420		
Db	361	GGTGATGGCAGGGGACCTGCTTTTCGCCCGGCTATCTTACTTTGAAAGGCTCCTCG	420		
Qy	421	GGAGGCCCTCTGTGTGCCCGCAGAGCATGTCGGTAGGCATATTCAGAGCGGGTATGC	480		
Db	421	GGAGGCCCTCTGTGTGCCCGCAGAGCATGTCGGTAGGCATATTCAGAGCGGGTATGC	480		
Qy	481	ACCGTGGAGTGGCTAAGGGGTGGACTTATCCCGTAGAGCTTTAGAGACAACCATG	540		
Db	481	ACCGTGGAGTGGCTAAGGGGTGGACTTATCCCGTAGAGCTTTAGAGACAACCATG	540		
Qy	541	AGGTCCCGGTGTTCTCAGACAACTCCTCCCGCAGCAGTCCCGCAGAGCTACCAAGTG	600		
Db	541	AGGTCCCGGTGTTCTCAGACAACTCCTCCCGCAGCAGTCCCGCAGAGCTACCAAGTG	600		
Qy	601	GCCACCTGCATGCTCCACCGGCGGGTAAGAGCACCAAGGTCCCGGCGCATACGCA	660		
Db	601	GCCACCTGCATGCTCCACCGGCGGGTAAGAGCACCAAGGTCCCGGCGCATACGCA	660		
Qy	661	GCTCAGGGCTACAGGTGCTGCTCAACCCCTCCGTTGCTGCAACAATGGGCTTTGGT	720		
Db	661	GCTCAGGGCTACAGGTGCTGCTCAACCCCTCCGTTGCTGCAACAATGGGCTTTGGT	720		
Qy	721	GCTTACATGTCCAAAGGCCCATGGATTGATCCTAAACATCAGGACTGGGGTGAGACAAT	780		
Db	721	GCTTACATGTCCAAAGGCCCATGGATTGATCCTAAACATCAGGACTGGGGTGAGACAAT	780		
Qy	781	ACTATGGGACCGCGATCAGTATTCACCTTAGCGGAAGTTCCTTCCGACGGGGTGT	840		
Db	781	ACTATGGGACCGCGATCAGTATTCACCTTAGCGGAAGTTCCTTCCGACGGGGTGT	840		
Qy	841	TCAGGGGGTCTTATGACATAATTTGTGACAGTGCCTCCACTCCAGGATGCACATCC	900		
Db	841	TCAGGGGGTCTTATGACATAATTTGTGACAGTGCCTCCACTCCAGGATGCACATCC	900		
Qy	901	ATCTTGGGCACTGGCACTGCTTGAACAAGCAGAGACCGGGGGGAGACTGACTGTG	960		
Db	901	ATCTTGGGCACTGGCACTGCTTGAACAAGCAGAGACCGGGGGGAGACTGACTGTG	960		
Qy	961	CTCGCCACCGCTACCCCTCCGGGCTCCGTCATCTGTGCCCATCTCTAACATCGAGGAGTT	1020		

961	CTCGCCACCGCTACCCCTCCGGGCTCCGTCATCTGTGCCCATCTCTAACATCGAGGAGTT	1020
1021	GCTCTGTCCACTACCGGAGAGATCCCTTTTATGGCAAGGCTATTCCTCTTGAAGCAAT	1080
1021	GCTCTGTCCACTACCGGAGAGATCCCTTTTATGGCAAGGCTATTCCTCTTGAAGCAAT	1080
1081	AAGGGGGGAGACATCTCTTCTGCACTCAAAAGAAAGTGCAGAGCTCGCGCA	1140
1081	AAGGGGGGAGACATCTCTCTTCTGCACTCAAAAGAAAGTGCAGAGCTCGCGCA	1140
1141	AAACTGGTCCGCTTGGCGCTCAATGCGTGCGCTTACTACCGCGCTTGTATGTCTCGTC	1200
1141	AAACTGGTCCGCTTGGCGCTCAATGCGTGCGCTTACTACCGCGCTTGTATGTCTCGTC	1200
1201	ATCCCGACAGTGGTGCAGTGTGCTGGCAACTGACGCCCTCATGACCGGCTTTACC	1260
1201	ATCCCGACAGTGGTGCAGTGTGCTGGCAACTGACGCCCTCATGACCGGCTTTACC	1260
1261	GGCGACTTCGATTCGGTGTAGACTGCACAACACGTGTGTACCCAGACAGTCGACTT	1320
1261	GGCGACTTCGATTCGGTGTAGACTGCACAACACGTGTGTACCCAGACAGTCGACTT	1320
1321	CTTGACCTTACCTTCAACATTGAGACATCAACGCTTCCCAGGATGCTCTCCGCTACT	1380
1321	CTTGACCTTACCTTCAACATTGAGACATCAACGCTTCCCAGGATGCTCTCCGCTACT	1380
1381	CAACGTCCGGGTAGGACTGGCAGAGGAAGCAGGCACTACAGATTTGTGGACCCGGG	1440
1381	CAACGTCCGGGTAGGACTGGCAGAGGAAGCAGGCACTACAGATTTGTGGACCCGGG	1440
1441	GAGCGTCTTCTGGCATGTTTGTGCTGTCTCTGCGAGTGTCTATGACGCGGTTGT	1500
1441	GAGCGTCTTCTGGCATGTTTGTGCTGTCTCTGCGAGTGTCTATGACGCGGTTGT	1500
1501	GCTTGTATGAGCTTACGCGCGCGAGACACAGTAGTAGGCTACGAGCATATGAACACC	1560
1501	GCTTGTATGAGCTTACGCGCGCGAGACACAGTAGTAGGCTACGAGCATATGAACACC	1560
1561	CCGGGACTTCCCGTGTGCCAAGACCATCTTGAATTTTGGAGGGCGCTTTTACGGGTCTC	1620
1561	CCGGGACTTCCCGTGTGCCAAGACCATCTTGAATTTTGGAGGGCGCTTTTACGGGTCTC	1620
1621	ACCCACATAGACGCCCATCTTCTATCCAGACAAAGCAGAGTGGGAAAACTTTCCCTAT	1680
1621	ACCCACATAGACGCCCATCTTCTATCCAGACAAAGCAGAGTGGGAAAACTTTCCCTAT	1680
1681	CTGGTAGCGTACCAAGCCACCGTGTGGCTAGAGCTCAAGCCCTCCCGCTGTGGGAC	1740
1681	CTGGTAGCGTACCAAGCCACCGTGTGGCTAGAGCTCAAGCCCTCCCGCTGTGGGAC	1740
1741	CAGATGTGAAGTGTGATCCGCTCTCAAGCCACCTCCATGGGCCAACACCTCTCTCTA	1800
1741	CAGATGTGAAGTGTGATCCGCTCTCAAGCCACCTCCATGGGCCAACACCTCTCTCTA	1800
1801	TATAGACTGGGCGCTGTCCAGAAATGAAGTCAACCTGACGACCCAGTCAACAAATATC	1860
1801	TATAGACTGGGCGCTGTCCAGAAATGAAGTCAACCTGACGACCCAGTCAACAAATATC	1860
1861	ATGACATGTATGTCCGCTGACCTTGGAGGTGCTCAAGGTACCTTGGGTGCTGTGGGCG	1920
1861	ATGACATGTATGTCCGCTGACCTTGGAGGTGCTCAAGGTACCTTGGGTGCTGTGGGCG	1920
1921	GTTCTGGCTGTTTGGCGGCTATTCCTATCCAGAGCTGCGTGGTTCATAGTAGGTAGG	1980
1921	GTTCTGGCTGTTTGGCGGCTATTCCTATCCAGAGCTGCGTGGTTCATAGTAGGTAGG	1980
1981	ATTGTCTTGTCCGAAAGCGGCAATCATATCCGACAGGGAAGTCTCTTACCGGGAGTTC	2040
1981	ATTGTCTTGTCCGAAAGCGGCAATCATATCCGACAGGGAAGTCTCTTACCGGGAGTTC	2040
2041	GATGAAATGGAAGTGTCTGA	2061







/product="non-structural protein 5b"  
/note="ORF10"

ORIGIN									
Query Match 90.0%; Score 1854.8; DB 14; Length 9610; Best Local Similarity 93.8%; Pred. No. 0; Matches 1931; Conservative 0; Mismatches 127; Indels 0; Gaps 0;									
QY	2	TGGCGCCTATCACGGCCTATGCCAGAGACAAGGGCCCTTTGGGATGCATATACCA	61	DB	4378	TCGCGACCGCTACCCCTCCGGGCTCGCTCACTGTGCCCCCATTTAAATCAAGAGGTTG	4437		
DB	3418	TGGGCCCATCACGGCTATGCCAGAGACAAGGGCCCTTTGGGATGCATATACCA	3477	QY	1022	CTCTGTCCACTACCGGAGAGATCCCTTTATGCGAAGGCTATTCCTCTTGAAGCAATTA	1081		
QY	62	GCTTGACCGCGCGGACAAAACAGAGTGGAGGTGAGGTTTCAGATCGTGTCAACTGCTG	121	DB	4438	CTCTGTCCACACCGGGGAGATTCCTTTTATGCGAAGGCTATTCCTCTTGAAGTATCA	4497		
DB	3478	GCCTGACCGCGCGGACAAAACAGAGTGGAGGTGAGGTTTCAGATCGTGTCAACTGCTG	3537	QY	1082	AGGGGGGAGACATCTCATCTTCTGCCACTCAAGAAGAAGTGCAGAGCTGCGCGCAA	1141		
QY	122	CCCAGACTTCTTGGCAACCTGCATTAACGGGGTGTGTGGACHTCTTCAATGGAGCGG	181	DB	4498	AGGGGGGAGACATCTCATCTTCTGCCACTCAAGAAGAAGTGCAGAGCTGCGCGCAA	4557		
DB	3538	CCCAGACTTCTTGGCAACCTGCATTAACGGGGTGTGTGGACHTCTTCAATGGAGCGG	3597	QY	1142	AACCTGTGCGGTGGGCGTCAATGCGGTGCTTACTACCGCGCTTGTGATGTGCGGTCA	1201		
QY	182	GAAACAAGGACCATTCGGTCACTAAGGTCCTGTATATCCAGATGTACACCAATGTGACC	241	DB	4558	AACCTGTGCGGTGGGCGTCAATGCGGTGCTTACTACCGCGCTTGTGATGTGCGGTCA	4617		
DB	3598	GAAACAAGGACCATTCGGTCACTAAGGTCCTGTATATCCAGATGTACACCAATGTGACC	3657	QY	1202	TCCCGACCAAGTGTGATGTGTGCTGTGCGCAACTGACGCCCTCATGACCGGCTTACCG	1261		
QY	242	AAGACCTCGTAGGTGCGCGCTCCCAAGGTGCGCGCTCATTAACCACTGCACTTGGC	301	DB	4618	TCCCGACCAAGTGTGATGTGTGCTGTGCGCAACTGACGCCCTCATGACCGGCTTACCG	4677		
DB	3658	AAGACCTCGTAGGTGCGCGCTCCCAAGGTGCGCGCTCATTAACCACTGCACTTGGC	3717	QY	1262	GGGACTTCGATTCGGGTGATAGACTGCAACACCTGTGTCAACAGAGTGTGCTCCCGTAC	1321		
QY	302	GCTCCTCGGACCTTTACTGTGTACGAGGACGCGCGATGCTATTCGTGCGCGGACGGG	361	DB	4678	GGGACTTCGATTCGGGTGATAGACTGCAACACCTGTGTCAACAGAGTGTGCTCCCGTAC	4737		
DB	3718	GCTCCTCGGACCTTTACTGTGTACGAGGACGCGCGATGCTATTCGTGCGCGGACGGG	3777	QY	1322	TTGACCTTACCTTACCATTTGAGACAATCACCGTTCCTCCCGAGGATGTGTCTCCCGTAC	1381		
QY	362	GTGATGCGAGGGGACGCTGCTTTCGCGCGCGCTATCTTACTTTGAAAGGCTCTCTCGG	421	DB	4738	TCGACCTTACTTTCACCATTTGAGACAAACACGCTTCCCGAGGATGTGTCTCCCGACCC	4797		
DB	3778	GTGATGCGAGGGGACGCTGCTTTCGCGCGCGCTATCTTACTTTGAAAGGCTCTCTCGG	3837	QY	1382	AACGTGCGGGTAGGACTGTGCGAGAGGAGCCAGGCACTTACAGATTTGTGGCAACGGGGG	1441		
QY	422	GAGCCCTCTGCTGTGCGCGCGAGACATCCGCTAGGCACTATTACAGACCGCGGTATGCA	481	DB	4798	AACGTGCGGGTAGGACTGTGCGAGGAGGAGCCAGGCACTTACAGATTTGTGGCAACGGGGG	4857		
DB	3838	GGGGCCCACTGCTGTGCGCTGCGGGAACACGCGGTGGGCAATTTACGGCGCGCGGTATGTA	3897	QY	1442	AGCGTCTTCTGCGCATTTTGTGACTGCTGCTCTGCTGCGAGTGTATGACGCGGGGTGTG	1501		
QY	482	CCGTGAGGTGGCTTAAGCGGTGGACTTTCATCCCGTATAGAGCTTAGAGACAACCATGA	541	DB	4858	AGCGCCCTTCCGGTATGTTGCACTCGCGCGTCTCTGCGAGTGTATGACGCGGGGTGTG	4917		
DB	3898	CCGAGAGGTGGCTTAAGCGGTGGAAATTTGCTCCCTGTAGAGAACCTTAGAGACAACCATGA	3957	QY	1502	CTTGTGTATGAGCTTACGCCCGCGGAGACACAGTTAGGCTACAGCATACATGAAACACCC	1561		
QY	542	GGTCCCGGTGTTCTCAGACAACTCTCTCCACAGCAGTGCCTCCAGAGCTACCAAGTGG	601	DB	4918	CTTGTGTATGAGCTACACCCCGCGGAGACACAGTTAGGCTACGAGCATACATGACACCC	4977		
DB	3958	GGTCCCGGTGTTTACAGACAACTCTCTCCACAGCAGTGCCTCCAGAGCTTCCAGGTGG	4017	QY	1562	CGGGACTTCCCGTGTGCGAAGCAATCTTGAATTTTGGAGGGCGTCTTTACGGGTCTCA	1621		
QY	602	CCACCTGCATGTCTCCACCGGAGCGGTAAAGACCAAGGTTCCTCCCGCGCATACGCGAG	661	DB	4978	CGGGACTTCCCGTGTGCGAAGCAATCTTGAATTTTGGAGGGCGTCTTTACCGGTCTCA	5037		
DB	4018	CCACCTGCATGTCTCCACCGGAGCGGTAAAGACCAAGGTTCCTCCCGCGCATACGCGAG	661	QY	1622	CCACATAGACGCCCACTTCTTATCCAGACAAAGCAGAGTGGGGAACCTTCCTCTATC	1681		
QY	662	CTCAGGGCTACAAGGTGCTGCTCAACCCCTCCGTGCTGCTGCAAAATGGGCTTTGGTG	721	DB	5038	CCCATATAGACGCTCACTTCTTATCCAGACAAAGCAGAGTGGGGAACCTTCTCTTACC	5097		
DB	4078	CTCAGGGTACAAGGTGCTGCTCAACCCCTCCGTGCTGCTGCAAAATGGGCTTTGGTG	4137	QY	1682	TGGTAGCGTACCAAGCCACCGTGTGCGCTTAGAGCTCAAGCCCCCTCCCGCTGCTGGGACC	1741		
QY	722	CTTACATGTCCAAGGCCATCGGATTCATCTTAACATCAGGACTGGGGTGAAGCAATTA	781	DB	5098	TGGTAGCGTACCAAGCCACCGTGTGCGCTTAGAGCCCAAGCCCCCTCCCGCTGCTGGGACC	5157		
DB	4138	CCTATATGTCCAAGGCCATCGGATTCATCTTAACATCAGGACTGGGGTGAAGCAATTA	4197	QY	1742	AGATGTGGAAGTGTGATTCGCTTCAAGCCCAACCTTCATGCGGCAACACCTCTGTCTAT	1801		
QY	782	CTACTGGCAGCCGATCACTGATTCACCTTACGGCAAGTTCTTTCGCGACGGCGGTGTT	841	DB	5158	AGATGTGGAAGTGTGATTCGCTTCAAGCCCAACCTTCATGCGGCAACACCTCTGTCTAT	5217		
DB	4198	CCACCGGACGCCCATCACTACTTCCACTTACGGCAAGTTCTTTCGCGACGGCGGTGCT	4257	QY	1802	ATAGACTCGGGCTGTCTCAGAAATGAAGTCAACCTGACGCCACCCAGTCAACAGTATATCA	1861		
QY	842	CAGGGGTGCTTATGACATAAATTTGTGACGAGTGCCACTCCACGGATGCACATATCA	901	DB	5218	ACAGACTGGGCGCTGTTTCAAGATGAAGTCAACCTTAACGCCAGTCAACCAATATATCA	5277		
DB	4258	CAGGGGTGCTTATGACATAAATTTGTGACGAGTGCCACTCCACGGATGCACATATCA	4317	QY	1862	TGACATGTATGTGCGCTGACCTTGGAGGTGCTCACAGGTACCTGGGTGCTCGTTGGCGCG	1921		
QY	902	TCTTGGGCAATGGCACTGTCTTTGACCAAGCAGAGACCGCGGGGCGAGACTGACTGTGC	961	DB	5278	TGACATGTATGTGCGCTGATCTGGAGATCGTCAAGGTACCTGGGTGCTCGTTGGCGCG	5337		
DB	4318	TCTTGGGCAATGGCACTGTCTTTGACCAAGCAGAGACCGCGGGGCGAGGCTGGTGTGC	4377	QY	1922	TTCCTGGCTGCTTTGGCGCGCTATTCCTTATCCAGGCTGTGGTGTATAGTAGGTAGGA	1981		
QY	962	TGCCACCGCTACCCCTCCGGGCTCCGTCACTGTGCCCCCATTCCTAACATCGAGGAGTTG	1021	DB	5338	TTCTGGCTGCTTTGGCGCGCTACTGTCTATCCAGGCTGTGGTGTATAGTAGGTAGGA	5397		
				QY	1982	TTGTCTTCTCGGAAAGCGCGCAATCATACCCGACAGGGAAGTCTCTACCGGAGTTTCG	2041		
				DB	5398	TTGTCTTCTCGGAAAGCGCGCAATCATACCCGACAGGGAAGTCTCTACCGGAGTTTCG	5457		
				QY	2042	ATGAAATGGAAGAGTGTCT 2059			



QY 782 CTACTGGCAGCCCGATCATCGTATTCACCTACGGCAAGTTCTTTCGCCGACGGCGGGTGT 841  
Db |||||  
QY 3923 CTACTGGCAGCCCGATCATCGTATTCACCTACGGCAAGTTCTTTCGCCGACGGCGGGTGT 3982  
Db |||||  
QY 842 CAGGGGTGCTTATGACATAATATTTGTGACGATGCGCACTCCACGGATGCAACATCCA 901  
Db |||||  
QY 3983 CAGGGGTGCTTATGACATAATATTTGTGACGATGCGCACTCCACGGATGCAACATCCA 4042  
Db |||||  
QY 902 TCTTGGGCATTTGGCAGTCTCTTGCACCAAGCAGAGACGGCGGGCGAGACTGCTGC 961  
Db |||||  
QY 4043 TCTTGGGCATTTGGCAGTCTCTTGCACCAAGCAGAGACTGCGGGGCTAGACTGCTGCTGC 4102  
Db |||||  
QY 962 TCGCCACCGCTACCCCTCCGGGCTCCGCTCACTGTGCCCCCATCTTAACATCGAGAGGTG 1021  
Db |||||  
QY 4103 TCGCCACCGCTACCCCTCCGGGCTCCATCATCTGTGCCCATCTTAACATCGAGAGGTG 4162  
Db |||||  
QY 1022 CTCTGTCACTACCGGAGAGATCCCTTTTATGCAAGGCTATTCCTCTTGAAGCAATTA 1081  
Db |||||  
QY 4163 CTCTGTCCACACCGGAGAGATCCCTTTTATGCAAGGCTATTCCTCTTGAAGCAATTA 4222  
Db |||||  
QY 1082 AGGGGGGAGACATCTCATCTTCTGCCACTCAAGAGAGAGTGCAGAGCTCGCCGCAA 1141  
Db |||||  
QY 4223 AGGGGGGAGACATCTCATCTTCTGCCACTCAAGAGAGAGTGCAGAGCTCGCCGCAA 4282  
Db |||||  
QY 1142 AACTGGTGGCTTGGGCTCAATCCGCTTACTACCGCGGCTTGTGTGCTGCTCA 1201  
Db |||||  
QY 4283 AACTGGTGGCTTGGGCTCAATCCGCTTACTACCGCGGCTTGTGTGCTGCTCA 4342  
Db |||||  
QY 1202 TCCGACAGTGTGAGTGTGTGCTGTGCACTGAGCCCTCATGACCGGCTTTACCG 1261  
Db |||||  
QY 4343 TCCGACAGTGTGAGTGTGTGCTGTGCACTGAGCCCTCATGACCGGCTTTACCG 4402  
Db |||||  
QY 1262 GCGACTTCGATGATGACTGCAACACGCTGTGTCACCCAGACAGTGCATTCAGCC 1321  
Db |||||  
QY 4403 GCGACTTCGATGATGACTGCAACACGCTGTGTCACCCAGACAGTGCATTCAGCC 4462  
Db |||||  
QY 1322 TTGACCTTACCTTCACTTGGAGCAATCACTGCTTCCCGAGTGTGCTTCCCGTACTC 1381  
Db |||||  
QY 4463 TTGACCTTACCTTCACTTGGAGCAATCACTGCTTCCCGAGTGTGCTTCCCGTACTC 4522  
Db |||||  
QY 1382 AAGTCCGGGTAGGCTGCGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1441  
Db |||||  
QY 4523 AAGTCCGGGTAGGCTGCGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 4582  
Db |||||  
QY 1442 AGCGTCTCTTGGCATGTTTGAATCTGCTCTCTGCGAGTGTGCTATGACCGGGTGTG 1501  
Db |||||  
QY 4583 AGCGCTCTTGGCATGTTTGAATCTGCTCTCTGCGAGTGTGCTATGACCGGGTGTG 4642  
Db |||||  
QY 1502 CTTGGTATGAGCTTACCGCCCGGAGACACAGTGTAGGCTACGAGCATATGAAACACCC 1561  
Db |||||  
QY 4643 CTTGGTATGAGCTTACCGCCCGGAGACACAGTGTAGGCTACGAGCATATGAAACACCC 4702  
Db |||||  
QY 1562 CGGAGCTTCCCGGTGCGAGAGCCATCTTGAATTTTGGGAGGCGCTTTTACGGTCTCA 1621  
Db |||||  
QY 4703 CGGAGCTTCCCGGTGCGAGAGCCATCTTGAATTTTGGGAGGCGCTTTTACGGTCTCA 4762  
Db |||||  
QY 1622 CCACATAGACGCCCACTTCTATCCAGACAAAGCAGAGTGGGGAACCTTCCCTATC 1681  
Db |||||  
QY 4763 CCACATAGACGCCCACTTCTATCCAGACAAAGCAGAGTGGGGAACCTTCCCTATC 4822  
Db |||||  
QY 1682 TGTAGGCTACCAAGCAGCGTGTGGCTAGAGCTCAAGCCCTTCCCGTGTGGGAC 1741  
Db |||||  
QY 4823 TGTAGGCTACCAAGCAGCGTGTGGCTAGAGCTCAAGCCCTTCCCGTGTGGGAC 4882  
Db |||||  
QY 1742 AGATGTGGAAGTCTTGTATCCGCTCAAGCCCACTTCCATGGCCCAACCTCTGCTAT 1801  
Db |||||  
QY 4883 AGATGTGGAAGTCTTGTATCCGCTCAAGCCCACTTCCATGGCCCAACCTCTGCTAT 4942  
Db |||||  
QY 1802 ATAGCTGGCGGTGTCCAGATGAAGTCACTGTGACGACCCAGTCAACCAAGTATATCA 1861  
Db |||||  
QY 4943 ACAGGCTGGGTGCGGTTCAGATGAAGTCACTGTGACGACCCAGTCAACCAAGTATATCA 5002  
Db |||||  
QY 1862 TGACATGTATGTGGCTGACCTGGAGGTGCTCACGAGTACCTGGGTGCTGCTGGCGG 1921  
Db |||||

Db 5003 TGACATGCATGTCGGCGACCTGGAGTGTGTCACGAGTACCTGGGTGCTCGTGGTGGCG 5062  
QY 1922 TTCTGGCTGCTTTGGCGCGTATTTGCTATCCACAGGCTCGCTGGTTCATAGTAGGTAGGA 1981  
Db |||||  
QY 5063 TTCTGGCTGCTTTGGCGCGTATTTGCTATCCACAGGCTCGCTGGTTCATATAGGACGGA 5122  
Db |||||  
QY 1982 TTCTCTTGTCCGGAAGCGCGGATCATATACCGACAGGGAAGTCTCTTACCGGAGTTCCG 2041  
Db |||||  
QY 5123 TTCTCTTGTCCGGAAGCGCGGAGTATACCGGACAGGAGGTCTCTTACCGAGAGTTCCG 5182  
Db |||||  
QY 2042 ATGAAATGGAAGAGTGTCT 2059  
Db |||||  
QY 5183 ATGAAATGGAAGAGTGTCT 5200  
Db |||||

RESULT 6  
AX164584  
LOCUS AX164584 6299 bp DNA linear PAT 22-JUN-2001  
DEFINITION Sequence 6 from Patent WO0138360.  
ACCESSION AX164584  
VERSION AX164584.1 GI:14545518  
KEYWORDS synthetic construct  
SOURCE synthetic construct  
ORGANISM other sequences; artificial sequences.  
REFERENCE 1  
AUTHORS Coit,D.C., Medina-Selby,A.C., Selby,M.C. and Houghton,M.C.  
TITLE Novel hcv non-structural polypeptide  
JOURNAL Patent: WO 0138360-A 6 31-MAY-2001;  
CHIRON CORPORATION (US)  
FEATURES  
source  
1..6299  
/organism="synthetic construct"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:32630"  
/note="pNS34a"  
1990..4050  
/note="unnamed protein product"  
/codon\_start=1  
/transl\_table=11  
/protein\_id="CAC42724.1"  
/db\_xref="GI:14545519"  
/translation="MAPITAYAAQOTRGLGCIITSLTGRDKNOVEGEVQIVSTAAQTFL  
LATCTGVCVTYVHAGRTETIASPKGVIOYMTAVDOLVGPASQGRSLTPTCTGSG  
SDLYLVTREADVIPRRGRDSRGLSPRISYLKSGSGGFLCPAGHAGVIFRAVC  
TRGAKAVDFIPVENLETTMRSPVTDNSPPVQSPQVAHLHAGTSGSKTKVPA  
YAAQGYKVLNLPVAAATLFGAYNSKAHGIDPNRTGRTITTPGSPITYGKFLA  
DGGSGGAYDIIICDECHSDATSLIGITVLDQAEATAGARLVATATPSSVTPH  
PNIIEVALSTTTEIPFYGKAIPLEVIKGRHLIFCHSKKDELAALVALGINAVAY  
YRGLDVSPITSGDVVVVATDALMTGTVDFDSDVDNCTCVTQTVDFSLDPTFTETI  
TLPDVSRTRRGTGRGKGIYRFVAPGRPSGMDSSVLCECYDAGCAWELTPA  
ETTVKRAYMNTPGLPVQDHLFEWEGVTGLTHIDAFLSQTKQSGENLPLVAYQA  
TVCARAQAPPSQDMKCLIRLPTLHGPTPLLYRLGAVQNEITLTHPVTKYIMTCM  
SADLEVTSTWLVGGVLAALAAAYCLSTGCVVIVGVVLSGKPAIIPREVLRYRFE  
MEEC"

ORIGIN  
Query Match 86.7%; Score 1787; DB 6; Length 6299;  
Best Local Similarity 91.7%; Pred. No. 0;  
Matches 1889; Conservative 0; Mismatches 170; Indels 0; Gaps 0;  
QY 1 ATGGCGCTATACAGCGCTATGCCAGCAGACAAGGGGCTTTTGGGATGATATACAC 60  
Db 1990 ATGGCGCCCATACAGCGCTACGCCAGCAGACAAGGGGCTCTCTAGGTGATATACAC 2049  
QY 61 AGCTTGACCGCGGACAAACCAAGGTGAGGAGGTTCAGATCGTGTCAACTGCT 120  
Db 2050 AGCCTAACTGCGCGGACAAACCAAGGTGAGGAGGTTCAGATCGTGTCAACTGCT 2109  
QY 121 GCCCAGACTTTCTTGGCAACCTGATTAACCGGGGTGTGTGGACTGTCTACATGGAGCC 180  
Db 2110 GCCCAAACTTTCTCTGGCAACCTGATCAATGGGGGTGTGTGGACTGTCTACACGGGGCC 2169







Db	2971	ATGAGATGGGAAGTGTCT	2988
RESULT 8	I09328		
LOCUS	I09328	5360 bp	DNA linear PAT 02-DEC-1994
DEFINITION	Sequence 8 from Patent WO 8904669.		
ACCESSION	I09328		
VERSION	I09328.1	GI:587963	
KEYWORDS	.		
SOURCE	Unknown.		
ORGANISM	Unknown.		
REFERENCE	1 (bases 1 to 5360)		
AUTHORS	Houghton,M., Choo,Q.-K. and Kuo,G.		
JOURNAL	Patent: WO 8904669-A 8 01-JUN-1989;		
FEATURES	Location/Qualifiers		
source	1..5360		
ORIGIN	/organism="unknown"		
	/mol_type="unassigned DNA"		
Query Match	86.7%;	Score 1786;	DB 6; Length 5360;
Best Local Similarity	91.7%;	Pred. No. 0;	
Matches 188;	Conservative 0;	Mismatches 170;	Indels 0; Gaps 0;
Qy	2	TGGCGCCCTATCAGCGCCTATGCCAGCAGACAAGGGGCCTTTTGGGATGCATAATCACCA	61
Db	931	TGGGCCCATCACGGCGTAGCCCGACGACGACAGGGGCCTCTTAGGGTGCAATACCA	990
Qy	62	GCTTGACCGCGGGGACAAAACACAGGTGAGGGTGAGGTTTCAGATCGTGTCAACTGCTG	121
Db	991	GCCTAACTGCGCGGGACAAAACCAAAGTGAGGGTGAGGTCAGATTGTGTCAACTGCTG	1050
Qy	122	CCAGACTTTCTTGGGCAACTGTCAATTAACGGGGTGTGTGGACTGTCTACCATTGGAGCCG	191
Db	1051	CCCAAACTTTCTGGCAACGTGCATCAATGGGGTGTGTGGACTGTCTACCAAGGGGCGG	1110
Qy	182	GAACAAGGACATTGCGTCACTAACGGTCTCTGTATCCAGATGTACACCAATGTGGACC	241
Db	1111	GAAGGAGACATCGGTCAACCAAGGGTCTGTCTATCCAGATGTATACCAATGTAGACC	1170
Qy	242	AAGACTCTGTAGGCTGGCCCGCTCCCCAAGGTGCCCGCTCATTAACACCAATGCACTTGGC	301
Db	1171	AAGACTTGTGGCTGGCCCGCTCCGCAAGGTAGCCGCTCATTTGACACCCCTGCACCTTGG	1230
Qy	302	GCTCTCGGACCTTTACTCGTCAAGGACAGCCGATGTCAATCTGTGCGCCGAGGGG	361
Db	1231	GCTCTCGGACCTTTACTCGTCAAGGACAGCCGATGTCAATCTCCGCTGCGCCGCGGG	1290
Qy	362	GTGATGCGAGGGCAGGCTGCTTTTGGCCCGGCTATCTCTTAAAGGCTCTCTCGG	421
Db	1291	GTGATGCGAGGGCAGGCTGCTGTGCCCCCGGCCATTTCTACTTTAAAGGCTCTCTCGG	1350
Qy	422	GAGGCCCTGTGCTGTGCCCGCAGGACATGCCGTAGGCATATTTACAGAGCCGCGGTATGCA	481
Db	1351	GGGGTCCGCTGTTGTGCCCCCGGGCAGCCGCTGGGCATATTTAGGGCCGCGGTGCA	1410
Qy	482	CCCGTGAGTGGCTTAAGCGGTGGAATTTCATCCCCTGTAGAGCTTAGAGACAACATGA	541
Db	1411	CCCGTGAGTGGCTTAAGCGGTGGAATTTATCCCTGTGGAGAACCTTAGAGACAACCATGA	1470
Qy	542	GGTCCCGGTGTTCTCAGACAATCTCTCCACACAGAGTCCCGAGAGCTACCAAGTGG	601
Db	1471	GGTCCCGGTGTTTACGGGTAATCTCTCTCCACAGTAGTGCCCCAGAGCTTCCAGGTGG	1530
Qy	602	CCCACTGTGATGTTCCCAACCGGAGCGGTAAAGACCAAGAAGTCCCGGCCGACATACGCAG	661
Db	1531	CTCACTCCATGCTCCACAGGAGCGCGCAAGACCAAGAGTCCCGGCTGCATATGCAG	1590
Qy	662	CTCAGGCTCAAGGTGCTGGTGTCAAAACCCCTCCGTTGTGTGCAAAATAGGGCTTTGGTG	721
Db	1591	CTCAGGCTCAAGGTGCTAGTACTCAACCCCTCTGTGTGTGCAACACTGGGCTTTGGTG	1650

Qy 1802 ATAGACTGGCGCTGTCAGAAATGAATCAACCTGACGACCCAGTCACCAAGTATATCA 1861  
Db 2731 ACAGACTGGCGCTGTTTCAAGATGAATCAACCTGACGACCCAGTCACCAATATATCA 2790  
Qy 1862 TGACATGATGTCGGCTGACCTGGAGTCTGTCAGAGTACCTGGGTGCTGTTGGCGCG 1921  
Db 2791 TGACATGATGTCGGCGGACCTGGAGTCTGTCAGAGTACCTGGGTGCTGTTGGCGCG 2850  
Qy 1922 TTCTGGCTGCTTTGGCGCGCTATTGCTTATCCACAGGCTCGTGTGTCATAGTAGTAGGA 1981  
Db 2851 TCCTGGCTGCTTTGGCGCGCTATTGCTTATCCACAGGCTCGTGTGTCATAGTAGTAGGA 2910  
Qy 1982 TTGTCTTGTCCGAAAGCCCGCAATCATACCCGACAGGGAAGTCTCTTACCGGAGTTG 2041  
Db 2911 TCGTCTTGTCCGGAAGCCCGCAATCATACCTGACAGGGAAGTCTCTTACCGAGTTG 2970  
Qy 2042 ATGAATGGAAGTGTCT 2059  
Db 2971 ATGAGATGGAAGTGTCT 2988

RESULT 9  
I06440  
LOCUS I06440 6785 bp DNA linear PAT 02-DEC-1994  
DEFINITION Sequence 54 from Patent EP 0318216.  
ACCESSION I06440  
VERSION I06440.1 GI:590312  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 6785)  
AUTHORS Houghton,M., Choo,Q.-L. and Kuo,G.  
TITLE Nanbv diagnostics and vaccines  
JOURNAL Patent: EP 0318216-A1 54 31-MAY-1989;  
FEATURES  
source 1..6785  
/organism="unknown"  
/mol\_type="unassigned DNA"

ORIGIN  
Query Match 86.7%; Score 1786; DB 6; Length 6785;  
Best Local Similarity 91.7%; Pred. No. 0;  
Matches 1888; Conservative 0; Mismatches 170; Indels 0; Gaps 0;  
Qy 2 TGGCGCTATCAGCGCTATGCCAGCAGACAAAGGGGCTTTGGGATGCATAATCACCA 61  
Db 1204 TGGCGCCATCAGCGGTACGCCAGCAGACAAAGGGGCTCTTAGGGTGCAATAATCACCA 1263  
Qy 62 GCTTGACCGCGCGGACAAACACAGGTGGAGGTGAGGTTTCAGATCGTGTCAACTGCTG 121  
Db 1264 GCCTAACTGCGCGGACAAACACAGGTGGAGGTGAGGTTTCAGATCGTGTCAACTGCTG 1323  
Qy 122 CCAGACTTTCTTGGCAACTGATTAACGGGTGTGTTGGACTGTCTACCATGGAGCGG 181  
Db 1324 CCCAAACCTTCTGGCAACTGATCAATGGGTGTGCTGGAAGTGTCTACCGCGGCGG 1383  
Qy 182 GRACAGGACATTTGCTGTCACCTAAGGTCCTGTTATCCAGATGTACCAATGTGACC 241  
Db 1384 GAACGAGGACATTCGGGTACCCCAAGGTCCTGTCATCCAGATGTATCAATGTAGACC 1443  
Qy 242 AAGACCTCGTAGGCTGGCGCTCCCAAGGTGCCCGCTCATTAACACCATGCACTTGGC 301  
Db 1444 AAGACCTTGTGGCTGGCGCTCCGCAAGGTAGCCGCTCATTTGACACCTGCACTTGG 1503  
Qy 302 GCTCTCGGACCTTTACCTGTCACAGGACCGCGATGTCATCTCTGTGCGCGACGGG 361  
Db 1504 GCTCTCGGACCTTTACCTGTCACAGGACCGCGATGTCATCTCTGTGCGCGCGGG 1563  
Qy 362 GTGATGGCAGGGCAGCTGCTTTCGCCCGGCTATCTCTTACTTGAAGGCTCTCTCGG 421  
Db 1564 GTGATGACGGGACGCTGCTGTCGCCCGGCGCCATTTCTTACTTGAAGGCTCTCTCGG 1623

Qy 422 GAGGCCCTCTGTGTGTCGCCCGCAGGACATGCGGTAGGCATATTTCAGAGCCGCGGTATGA 481  
Db 1624 GGGTCCGCTGTGTGTCGCCCGCGGGGACGCGGTGCGCATATTTAGGGCCGCGGTGTGA 1683  
Qy 482 CCCGTGGAGTGGCTAAGGGCGGTGGACTTCATCCCGTAGAGACTTTAGAGACAAACATGA 541  
Db 1684 CCCGTGGAGTGGCTAAGGGCGGTGGACTTCATCCCTGTGGAGAACCTTAGAGACAAACATGA 1743  
Qy 542 GGTCCCGGCTGTCTCAGACAACTCCTCCACACGACGAGTGGCCAGAGCTACCAAGTGG 601  
Db 1744 GGTCCCGGCTGTTCACGATAAATCTCTCTCCACAGTAGTGGCCAGAGCTTCCAGGTGG 1803  
Qy 602 CCACCTGTCATGCTCCACCGCGCAGCGGTAAAGACACCAAGGTTCGCCGCGCATACGACG 661  
Db 1804 CTCACCTCCATGCTCCACAGGACGCGCAAGACACCAAGGTTCGCCGCTGCATATGACG 1863  
Qy 662 CTCAGGGCTACAAAGTGTGCTGCTCAACCCCTCCGTTCTGCTGCAACAATGGGGCTTTGGTG 721  
Db 1864 CTCAGGGCTATAAGGTGCTAGTACTCAACCCCTCTGTGCTGCAACACTGGGCTTTGGTG 1923  
Qy 722 CTTACATGTCCTAAGGCCCATGGATTGATCTTAAATCAGGACTGGGTGAGACAAATTA 781  
Db 1924 CTTACATGTCCTAAGGCTCATGGATCGATCTTAAATCAGGACCGGGGTGAGACAAATTA 1983  
Qy 782 CTACTGGCAGCCCGATCAGTATTCACCTACGGCAAGTTCTTTCGCCGACGCGGTGT 841  
Db 1984 CCACTGGCAGCCCATCAGTACTCCACCTACGGCAAGTTCTTTCGCCGACGCGGTGT 2043  
Qy 842 CAGGGGCTCTTATGACATAAATTTGTGACGAGTGCCACTCCACGGATGCAACATCCA 901  
Db 2044 CGGGGGCGCTTATGACATAAATTTGTGACGAGTGCCACTCCACGGATGCAACATCCA 2103  
Qy 902 TCTTGGCAATTCGCACTGCTTTGACAGCAGAGACGCGGGGGCGGACGACTGACTGTGC 961  
Db 2104 TCTTGGCAATTCGCACTGCTTTGACAGCAGAGACTGCGGGGGCGGACGACTGACTGTGC 2163  
Qy 962 TCGCCACGCTACCCCTCCGGGCTCCGTCACTGTGCCCATCTTAACTCGAGGAGTTG 1021  
Db 2164 TCGCCACGCGCACCCCTCCGGGCTCCGTCACTGTGCCCATCTTAACTCGAGGAGTTG 2223  
Qy 1022 CTCTGTCCACTACCGGAGAGATCCCTTTTATGGCAAGGCTATTCCCTTTGAAGCAATTA 1081  
Db 2224 CTCTGTCCACCGGAGAGATCCCTTTTATGGCAAGGCTATTCCCTTTGAAGCAATTA 2283  
Qy 1082 AGGGGGGAGACATCTCATCTTCTGCACTCAAGAAAGAGTGCGACGAGTCTCGCGCAA 1141  
Db 2284 AGGGGGGAGACATCTCATCTTCTGCACTCAAGAAAGAGTGCGACGAGTCTCGCGCAA 2343  
Qy 1142 AACTGTCGCGTTGGGCGTCAATGCCGTGGCTTACTACCGCGGCTTGTGTCCGTCA 1201  
Db 2344 AGCTGTCGCAATGGGCATCAATGCCGTGGCTTACTACCGGCTTGTGTCCGTCA 2403  
Qy 1202 TCCGACCATGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCG 1261  
Db 2404 TCCGACCATGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCG 2463  
Qy 1262 GGCATTCGATTCGGGTAGTAGTCAACACCTGTGTCAACCCAGACAGTCTGACTTCAGCC 1321  
Db 2464 GGCATTCGATTCGGGTAGTAGTCAACACCTGTGTCAACCCAGACAGTCTGACTTCAGCC 2523  
Qy 1322 TTGACCTACCTTACCAATGAGACAAATCACGCTTCCCGAGGATGCTGTCTCCCGTACTC 1381  
Db 2524 TTGACCTACCTTACCAATGAGACAAATCACGCTTCCCGAGGATGCTGTCTCCCGTACTC 2583  
Qy 1382 AACGTGCGGTTAGGACTGTCAGAGGAGCCAGGATCTTACAGATTTGTGGACCCGGGG 1441  
Db 2584 AACGTGCGGTTAGGACTGTCAGAGGAGCCAGGATCTTACAGATTTGTGGACCCGGGG 2643  
Qy 1442 AGCGTCTTCTTGGCATGTTTGTGCTGCTCTCTGCGAGTGTCTATGACGCGGTTGTG 1501  
Db 2644 AGCGTCTTCTTGGCATGTTTGTGCTGCTCTCTGCGAGTGTCTATGACGAGGCTGTG 2703  
Qy 1502 CTTGTATAGCTTACGCCCGCGGAGACCAAGTTAGGCTACGAGCATACATGAACCCC 1561

```
Db 2704 CTTGGTATAGCTCACCCCGCCGAGACTACAGTTAGGCTACGAGCTACATGAACACCC 2763
Qy 1562 CGGACTTCCCGTGTGCCAAGACCATCTTGAATTTTGGAGGGCGCTCTTTACGGGTCTCA 1621
Db 2764 CGGGCTTCCCGTGTGCCAAGACCATCTTGAATTTTGGAGGGCGCTCTTTACAGCCCTCA 2823
Qy 1622 CCACATAGAGCCCACTTCTATCCAGACAAAGCAGAGTGGGAAACCTTCCCTATC 1681
Db 2824 CTCATATAGATGCCCACTTTCTATCCAGACAAAGCAGAGTGGGAGAACTTCTTACC 2883
Qy 1682 TGGTAGGTTACCAAGCCACCGTGTGCGCTAGAGCTCAAGCCCTCCCGCTCGTGGGACC 1741
Db 2884 TGGTAGGTTACCAAGCCACCGTGTGCGCTAGGCTCAAGCCCTCCCGCTCGTGGGACC 2943
Qy 1742 AGATGTGGAAAGTCTTGATCTCGTCTCAAGCCCACTCCATGGGCCCAACACCTCTGCTAT 1801
Db 2944 AGATGTGGAAAGTGTGATCTCGTCTCAAGCCCACTCCATGGGCCCAACACCTCTGCTAT 3003
Qy 1802 ATAGACTGGGGCTGTCCAGAAATGAAGTCAACCTTGACGCCACCCAGTCAACCAAGTATATCA 1861
Db 3004 ACAGACTGGGGCTGTCCAGAAATGAAGTCAACCTTGACGCCACCCAGTCAACCAATATATCA 3063
Qy 1862 TGACATGTATCTGGCTGACCTGAGTGTGCTACAGAGTACCTGGGTGCTGTTGGCGGCG 1921
Db 3064 TGACATGCATGTGCGGCCGACCTGGAGTGTGCTACAGGACCTGGGTGCTGTTGGCGGCG 3123
Qy 1922 TTCTGGCTGTCTGGCGCGGTATGCTTATCCACAGCTGCTGCTCATAGTAGTAGGA 1981
Db 3124 TCCTGGCTGTCTGGCGCGGTATGCTTATCCACAGCTGCTGCTCATAGTAGTAGGA 3183
Qy 1982 TTGTCTTGTCCGAAAGCCGGAATCATATACCGACAGGGAAGTCTCTACCGGAGTTCG 2041
Db 3184 TCGTCTTGTCCGGAAGCCGGAATCATATACCTGACAGGGAAGTCTCTACCGAGAGTTCG 3243
Qy 2042 ATGAAATGGAAGTGTCT 2059
Db 3244 ATGAGATGGAAGTGTCT 3261

RESULT 10
ARL18696
LOCUS ARL18696 7310 bp DNA linear PAT 16-MAY-2001
DEFINITION Sequence 74 from patent US 6150087.
ACCESSION ARL18696
VERSION ARL18696.1 GI:14100606
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 7310)
AUTHORS Chien,D.Y.
TITLE NANBV diagnostics and vaccines
JOURNAL Patent: US 6150087-A 74 21-NOV-2000;
FEATURES
    Location/Qualifiers
        1..7310
            /organism="unknown"
            /mol_type="unassigned DNA"

ORIGIN

Query Match 86.7%; Score 1786; DB 6; Length 7310;
Best Local Similarity 91.7%; Pred. No. 0;
Matches 1888; Conservative 0; Mismatches 170; Indels 0; Gaps 0;

Qy 2 TGGGGCTATACGGCTATGCCCAGCAGACAGAGGGGCCCTTTGGGATGATATATACCA 61
Db 1729 TGGGCCCATACAGCGGTACGCCAGCAGACAGAGGGGCCCTCTAGGGTGCATATACCA 1788
Qy 62 GCTTGACCGCGCGACAAACACAGGTGGAGGTGAGGTTTCAGATCGTGTCAACTGCTG 121
Db 1789 GCCTAACTGGCGCGGACAAACACAGTGGAGGTGAGGTCAGATTTGTCTAACTGCTG 1848
Qy 122 CCCAGACTTTTCTGGCAACCTGCATTAACGGGGTGTGTGGACTGTCTACCATGGAGCGG 181
```

```
Db 1849 CCACAACTTCTCCGGCAACGTCATCAATGGGGTGTCTGGACTGTCTACACGGGCGG 1908
Qy 182 GAACAAGGACCATTTGGCTCACCTTAAGGTCTCTGTATCCAGATGTACACCAATGTGACC 241
Db 1909 GAACGAGGACCATTCGGCTCACCCAGGTCTCTGTATCCAGATGTATACCAATGTAGACC 1968
Qy 242 AAGACCTCTGTAGGTGCGCCGCTCCCAAGGTGCGGCTCATTAACACCATGTACTTGGC 301
Db 1969 AAGACCTTGTGGCTGGCCGCTCGCAAGGTAGCGCTCATTTGACACACCTCTGCACTTGGC 2028
Qy 302 GCTCTCTGGACCTTTACTTGGTCAAGAGGACGCGGATGTCTCTTGTGGCCCGACGGG 361
Db 2029 GCTCTCTGGACCTTTACTTGGTCAAGAGGACGCGGATGTCTCTTCCGCTGCGCGCGGG 2088
Qy 362 GTGATGACGAGGCGAGCTCTTTTGGCCCGGCTGTCTGCGCCCGGCCATTTCTTCTTGAAGGCTCTCTCGG 421
Db 2089 GTGATGACGAGGCGAGCTGTCTGCGCCCGGCCATTTCTTCTTGAAGGCTCTCTCGG 2148
Qy 422 GAGGCCCTCTGTGTGCCCCGCGAGGACATGCGGTAGGACATATTTCAGAGCCGCGGTATGCA 481
Db 2149 GGGGTCCGCTGTGTGCCCCGCGGCGACGCGGTGGGSCATATTTAGGGCCGCGGTGTGCA 2208
Qy 482 CCCGTGAGGTGGCTAAGCGGTGGACTTTCATCCCGGTAGAGAGCTTAGAGACAACCATGA 541
Db 2209 CCCGTGAGGTGGCTAAGCGGTGGACTTTCATCCCGGTAGAGAGCTTAGAGACAACCATGA 2268
Qy 542 GGTCCCCGCGTGTCTCAGACAACTCTCTCCACACAGCAGTCCCGCAGAGTACCAGTGG 601
Db 2269 GGTCCCCGCGTGTCTCAGACAACTCTCTCCACAGTGTGCCCCAGAGCTTCCAGGTGG 2328
Qy 602 CCACCTCTGATGTCTCCACCGCGAGCGGTAAAGACACAAAGTCTCCGCGCGCATACGACG 661
Db 2329 CTCACTCTCATGTCTCCACAGGCGAGCGGCAAAAGACCAAGGTCCCGGCTGCATATGACG 2388
Qy 662 CTGAGGCTCAAGAGTGTCTGCTCAACCCCTCGGTGTCTGCTGCAACATGGGCTTTGGTG 721
Db 2389 CTCAGGGCTATAAGAGTGTAGTACTCAACCCCTCTGTGTCTGCAACACTGGGGCTTTGGTG 2448
Qy 722 CTTTACATGTCCAAGGCCCATGGGATTTGATCTTAACATCAGGACTGGGGTGGAGCAATTA 781
Db 2449 CTTTACATGTCCAAGGCTCATGGGATCGATCTTAACATCAGGACCGGGGTGAGAACAAATTA 2508
Qy 782 CTACTGGCAGCCGATCACGTATTTCACCTACGCGCAAGTCTCTTGGCAGCGGGGTGTT 841
Db 2509 CCACTGGCAGCCCATCACGTACTCCACCTACGCGCAAGTCTCTTGGCAGCGGGGTGCT 2568
Qy 842 CAGGGGTGCTTATGACATTAATTTGTGACAGTGCCTCAACCGGATGCAACATCCA 901
Db 2569 CGGGGGCGGCTTATGACATTAATTTGTGACAGTGCCTCAACCGGATGCCACATCCA 2628
Qy 902 TCTTGGGCTATGGCACTGTCTTGACCAAGCAGAGACCGCGGGGGGAGACTGACTGTGC 961
Db 2629 TCTTGGGCTATGGCACTGTCTTGACCAAGCAGAGACTCGGGGGGGGAGACTGTTGTGC 2688
Qy 962 TCGCCACCGCTACCCCTCCGGGCTCGGTCACTGTGCCCCCATCTCTAAACATCGAGAGGTG 1021
Db 2689 TCGCCACCGCCACCCCTCCGGGCTCGGTCACTGTGCCCCCATCTCTAAACATCGAGGAGTGTG 2748
Qy 1022 CTCTGTCCACTACCGGAGAGATCCCTTTTATGGCAAGGCTATTCCTCTTGAAGCAATTA 1081
Db 2749 CTCTGTCCACCAACCGGAGAGATCCCTTTTATGGCAAGGCTATTCCTCTTGAAGCAATTA 2808
Qy 1082 AGGGGGGAGACATCTCATCTTCTGCACTCAAGAGAGTGGAGAGTGGAGAGTGGCCCAA 1141
Db 2809 AGGGGGGAGACATCTCATCTTCTGCTATTCAGAGAGAGTGGAGAGTGGAGAGTGGCCCAA 2868
Qy 1142 AACTGTGCGGTTGGGCGTCAATGCGGTGCTTACTACCGCGGCTTGTGTTGTGCTGCA 1201
Db 2869 AGCTGTGCTATTTGGGCACTCAATGCGGTGCTTACTACCGCGGCTTGTGAGGTGCTGCTGCA 2928
Qy 1202 TCCGACACAGTGTGTGCTGTGTGCGCAACTGACGCGGCTCATGACCGGCTTTTACCG 1261
```

Db	2929	TCCGACAGCGGCGATGTTGTCGTGCGTGGCAACCGATGCCCTCATGACGCGCTATACCG	2988
Qy	1262	GCAGCTTCGATTCGGTATAGACTGCAACACGCTGTGTCACCCAGACAGTGCATTCAGCC	1321
Db	2989	CGGACTTCGACTCGGTGATAGACTGCAATACGTTGTGTCACCCAGACAGTGCATTCAGCC	3048
Qy	1322	TTGACCTTACCTTACCAATTTGAGACATACGCTTCCCAAGGATGCTGTCTCCGTAATC	1381
Db	3049	TTGACCTTACCTTACCAATTTGAGACATACGCTTCCCAAGGATGCTGTCTCCGCACTC	3108
Qy	1382	AACGTCGGGTAGGACTGGCAGAGGGGAAGCAGGATCTACAGATTTGTGGCAGCGGGG	1441
Db	3109	AACGTCGGGGCAGGACTGGCAGAGGGGAAGCAGGATCTACAGATTTGTGGCAACCGGGG	3168
Qy	1442	AGCGTCCTTCGGCATGTTGACTCGTCTGCTCTCGAGTCTATGACGCGGTTGTG	1501
Db	3169	AGCGCCCTTCGGCATGTTGACTCGTCTGCTCTCGAGTCTATGACGCGGTTGTG	3228
Qy	1502	CTTGGTATGAGCTTACGCCCGCGAGACCAAGTTAGGCTACGAGCATATCAACACCC	1561
Db	3229	CTTGGTATGAGCTACGCCCGCGAGACTACAGTTAGGCTACGAGCTATCAACACCC	3288
Qy	1562	CGGACTTCCCGTGTGCCAAGACCACTTGAATTTTGGAGGGCGCTTTTACCGGCTCA	1621
Db	3289	CGGGGCTTCCCGTGTGCCAAGACCACTTGAATTTTGGAGGGCGCTTTTACAGGGCTCA	3348
Qy	1622	CCACATAGACGCCACTTCCCTATCCAGACAAGCAGAGTGGGGAAACCTTCCCTATC	1681
Db	3349	CTCATATAGATGCCACTTCTATCCAGACAAGCAGAGTGGGGAAACCTTCCCTATC	3408
Qy	1682	TGATAGCGTACCAAGCACCGTGTGGCTAGAGCTCAAGCCCTCCCGCTCGTGGGACC	1741
Db	3409	TGATAGCGTACCAAGCACCGTGTGGCTAGAGCTCAAGCCCTCCCGCTCGTGGGACC	3468
Qy	1742	AGATGGAAGTGCTGATCCGCTCMAAGCCACCTCCATGAGGGCAACACTCTGCTAT	1801
Db	3469	AGATGGAAGTGCTGATCCGCTCMAAGCCACCTCCATGAGGGCAACACTCTGCTAT	3528
Qy	1802	ATAGACTGGCGCTGTCCAGAAAGTCAAGCTGAGCTCAAGCCCTCCCGCTCGTGGGACC	1861
Db	3529	ACAGACTGGCGCTGTCCAGAAAGTCAAGCTGAGCTCAAGCCCTCCCGCTCGTGGGACC	3588
Qy	1862	TGACATGATGTGCGGTGACCTGGAGTGTGCTACGAGTACCTGGGTGCTGTTGGCGCG	1921
Db	3589	TGACATGATGTGCGCGACCTGGAGTGTGCTACGAGTACCTGGGTGCTGTTGGCGCG	3648
Qy	1922	TTCTGGCTGCTTTGGCGCGTATGCTATCCAGCTGCGTGGTCTATAGTAGGAGGA	1981
Db	3649	TTCTGGCTGCTTTGGCGCGTATGCTATCCAGCTGCGTGGTCTATAGTAGGAGGAG	3708
Qy	1982	TTGCTTTGTCGGAAAGCGGGCAATCATACCGACAGGGAAGTCTCTACCGGGAGTTGG	2041
Db	3709	TTGCTTTGTCGGGAAGCGGGCAATCATACCTGACAGGGAAGTCTCTACCGAGAGTTGG	3768
Qy	2042	ATGAAATGGAAGTGTCT 2059	
Db	3769	ATGAGATGGAAGTGTCT 3786	

RESULT 11  
109331

LOCUS 109331 7310 bp DNA linear PAT 02-DEC-1994  
DEFINITION Sequence 15 from Patent WO 8904669.  
ACCESSION 109331  
VERSION 109331.1 GI:587966

KEYWORDS Unknown.  
SOURCE Unknown.

ORGANISM Unclassified.

REFERENCE 1 (bases 1 to 7310)  
AUTHORS Houghton, M., Choo, Q.-K. and Kuo, G.  
JOURNAL Patent: WO 8904669-A 15 01-JUN-1989;  
FEATURES Location/Qualifiers

source	1. .7310 /organism="unknown" /mol_type="unassigned DNA"
ORIGIN	
Query Match	86.7%; Score 1786; DB 6; Length 7310;
Best Local Similarity	91.7%; Pred. No. 0;
Matches 1888; Conservative	0; Mismatches 170; Indels 0; Gaps 0;
Qy	2 TGGCGCTATACGCGCTATGCCAGACAGCAAGGGGCCCTTTGGGATGATATACCA 61
Db	1729 TGGCGCCATACGCGGTACGCCAGACAGAGGGGCCCTCTAGGGTGCATAATCACA 1788
Qy	62 GCTTGACCGCGCGGGAACAAAACAGGTGAGGTTCAGATTCAGATTCGATTCG 121
Db	1789 GCTTAATGCGCGGGAACAAAACAGTGGAGGTGAGTCCAGATTTGTCAACTGCTG 1848
Qy	122 CCAGACTTTCTGGCAACTGCAATTAACGGGGTGTGTGGAGTGTCTACCATGAGCGG 181
Db	1849 CCACAACTTCTGGCAACGTGCATCAATGGGGTGTGTGGACTGTCTACACGGGGCG 1908
Qy	182 GAAACAGGACCAATGCGTCACTAAGGTCTCTGTTATCCAGATGTACCAATGTGGAC 241
Db	1909 GAAACAGGACCAATGCGTCACTAAGGTCTCTGTTATCCAGATGTATACCAATGTAGAC 1968
Qy	242 AAGACCTCGTAGCGTGGCGCGCTCCCAAGGTGCCGCTCATTAACACCATGCACCTGCG 301
Db	1969 AAGACCTGTGGGTGGCGCGCTCCCAAGGTAGCGCTCATTTGACACCTGCACCTTGG 2028
Qy	302 GCTCTCGGACCTTTACCTGTTGACAGGACGCGGATGTCATTCCTGTGCGCGGACGG 361
Db	2029 GCTCTCGGACCTTTACCTGTTGACAGGACGCGGATGTCATTCCTGTGCGCGGACGG 2088
Qy	362 GTGATGGAGGGGACGCTGCTTTGCGCGCGGCTATCTCTTACTTGAAGGCTCTCGG 421
Db	2089 GTGATAGAGGGGACGCTGCTGTCGCGCGGCGCCATTTCTTACTTGAAGGCTCTCGG 2148
Qy	422 GAGGCCCTCTGCTGTGCGCGGACGACATGCGGTAGGACATTTACAGAGCGCGGTATGA 481
Db	2149 GGGGTCCGCTGTGTGCGCGGCGGACGCGGTGGGCATATTTAGGCGCGGCTGTGCA 2208
Qy	482 CCCGTGGAGTGTCTAAGCGGTGGAATTCATCCCGTAGAGAGCTTAGAGACAAACATCA 541
Db	2209 CCCGTGGAGTGTCTAAGCGGTGGAATTCATCCCGTGGAGAACCTTAGAGACAAACATCA 2268
Qy	542 GGTCCCGGTGTTCTCAGACAACTCGTCCCAACGACGAGTGCCTCCAGAGCTACCAAGTGG 601
Db	2269 GGTCCCGGTGTTCTCAGGATAACTCTCTCCACCGAGTAGTGCCTCCAGAGCTTCCAGGTGG 2328
Qy	602 CCCACCTGCATGCTCCCAACCGGTAAGAGCAACAAAGGTCCCGGCGCATACGCGAG 661
Db	2329 CTCACCTCATGCTCCCAACGAGCGGGCAAGACCAAGGTCCCGGCTGCATATGCA 2388
Qy	662 CTCAGGGCTACAAGGTGCTGTCTCAACCCCTCCGTTGCTGCTCAACAAATGGGCTTTGGTG 721
Db	2389 CTCAGGGCTATAAGGTGCTGTCTCAACCCCTCTGTTGCTCAACACACCTGGGCTTTGGTG 2448
Qy	722 CTTACATGTCGAAGGCCCATGGGATGTCTTAAATCATAGGACTGGGTGAGGACAAATTA 781
Db	2449 CTTACATGTCGAAGGCTCATGGGATCGATCTTAAATCATAGGACCGGGGTGAGGACAAATTA 2508
Qy	782 CTACTGGAGCGCCGATCACGTATTCACCTACGGCAAGTTCCTTGGCGACGGGGTGT 841
Db	2509 CCACCTGGAGCGCCCATCAGTACTCCACTACGGCAAGTTCCTTGGCGACGGGGTGT 2568
Qy	842 CAGGGGTGCTTATGACATAATAATTTGTGACGAGTGCACCTCCACGGATGCAACATCA 901
Db	2569 CCGGGGGGCTTATGACATAATAATTTGTGACGAGTGCACCTCCACGGATGCCATCA 2628
Qy	902 TCTTGGGCAATGGCACTGTCTTGACCAAGCAGACCGCGGGGGGAGCTGACTGTGCG 961
Db	2629 TCTTGGGCAATGGCACTGTCTTGACCAAGCAGACTGCGGGGGGCGAGCTGGTGTGCG 2688

962 TCGCCACCGGTACCCCTCCGGGCTCGCTCACTGTGCCCCCATCTTAAATCATGAGAGGTTG 1021  
 Db TCGCCACCGCACCCCTCCGGGCTCGCTCACTGTGCCCCCATCTTAAATCATGAGAGGTTG 2748  
 1022 CTCTGTCCATACCGGAGAGATCCCTTTTATGCAAGGCTATTCCCTTGAACCAATTA 1081  
 Db CTCTGTCCATACCGGAGAGATCCCTTTTATGCAAGGCTATTCCCTTGAACCAATTA 2808  
 1082 AGGGGGGAGACATCTCTCTTCTGCTGCAATCAAGAGAGAGTGCAGAGCTCGCCGCAA 1141  
 Db AGGGGGGAGACATCTCTCTTCTGCTGCAATCAAGAGAGAGTGCAGAGCTCGCCGCAA 2868  
 1142 AACTGTGCGCTTGGGCGTCAATCGCTGCTTACTACCGGCGCTTGTGATGTGCTGCTCA 1201  
 Db AACTGTGCGCTTGGGCGTCAATCGCTGCTTACTACCGGCGCTTGTGATGTGCTGCTCA 2928  
 1202 TCCGACGAGTGTGAGTGTGCTGCTGCGCAACTGACGCGCTCATGACGCGCTTTACCG 1261  
 Db TCCGACGAGTGTGAGTGTGCTGCTGCGCAACTGACGCGCTCATGACGCGCTTTACCG 2988  
 1262 GCGACTTTCGATTCGCTGATGACTGCAACACGCTGTGTCAACGACAGCTGCTTACGCC 1321  
 Db GCGACTTTCGATTCGCTGATGACTGCAACACGCTGTGTCAACGACAGCTGCTTACGCC 3048  
 1322 TTGACCTTACCTTACCATTTGAGCAATCAAGCTTCCCGAGGATGCTGTCTCCGCTACTC 1381  
 Db TTGACCTTACCTTACCATTTGAGCAATCAAGCTTCCCGAGGATGCTGTCTCCCGCACTC 3108  
 1382 AACGTCGGGTGAGGACTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1441  
 Db AACGTCGGGTGAGGACTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3168  
 1442 AGCGTCTCTTGGCATGTTTGAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1501  
 Db AGCGTCTCTTGGCATGTTTGAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3228  
 1502 CTTGCTGATGAGCTTACCGCGCGGAGACGACGATGAGCTGAGGCTACGAGCATATCA 1561  
 Db CTTGCTGATGAGCTTACCGCGCGGAGACGACGATGAGCTGAGGCTACGAGCATATCA 3288  
 1562 CGGAGCTTCCGCTGTCGACAGACCATCTGAAATTTTGGGAGGCGCTTTACGCGCTCA 1621  
 Db CGGAGCTTCCGCTGTCGACAGACCATCTGAAATTTTGGGAGGCGCTTTACGCGCTCA 3348  
 1622 CCCACATAGACGCGCTTCTTATCCAGACAAAGCAGAGTGCGGAGAACTTCCCTTATC 1681  
 Db CCCACATAGACGCGCTTCTTATCCAGACAAAGCAGAGTGCGGAGAACTTCCCTTATC 3408  
 1682 TGGTAGGCTACCAAGCAGCGTGTGCGCTAGAGCTCAAGCCCTTCCCGCTGCTGGGACC 1741  
 Db TGGTAGGCTACCAAGCAGCGTGTGCGCTAGAGCTCAAGCCCTTCCCGCTGCTGGGACC 3468  
 1742 AGATGTGGAAGTCTTGCATCGCTTCAAGCCCGCTTCCATGGGCGGAGGAGGAGGAGG 1801  
 Db AGATGTGGAAGTCTTGCATCGCTTCAAGCCCGCTTCCATGGGCGGAGGAGGAGGAGG 3528  
 1802 ATAGACTGGGCGCTTGCATCGGAGTCAAGCTTCAAGCCCGCTTCCATGGGCGGAGGAG 1861  
 Db ATAGACTGGGCGCTTGCATCGGAGTCAAGCTTCAAGCCCGCTTCCATGGGCGGAGGAG 3588  
 1862 TGACATGATGTGCGGCTGAGCTGAGGCTGCTGAGGAGTCTGAGGAGTCTGAGGAGGAG 1921  
 Db TGACATGATGTGCGGCTGAGCTGAGGCTGCTGAGGAGTCTGAGGAGTCTGAGGAGGAG 3648  
 1922 TTCTGGCTGCTTGGCGCGCTTATGCTTATCCAGAGCTGCTGCTGCTGCTGCTGCTGCTG 1981  
 Db TTCTGGCTGCTTGGCGCGCTTATGCTTATCCAGAGCTGCTGCTGCTGCTGCTGCTGCTG 3708  
 1982 TTGCTTGTGCGGAAAGCGGCAATCATACCGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2041  
 Db TTGCTTGTGCGGAAAGCGGCAATCATACCGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3768  
 2042 ATGAATGGAAGAGTGCT 2059

||||| 3769 ATGAGATGGAAGAGTGCT 3786  
 RESULT 12  
 HPCPOLYP  
 LOCUS  
 DEFINITION Hepatitis C virus polyprotein gene, partial cds.  
 ACCESSION M32084  
 VERSION M32084.1 GI:329875  
 KEYWORDS polyprotein.  
 SOURCE Hepatitis C virus  
 ORGANISM Hepatitis C virus  
 Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae; Hepacivirus.  
 1 (bases 1 to 7310)  
 Choo, Q.-L., Richman, K. and Han, J.  
 The nucleotide sequence of the Hepatitis C viral genome  
 Unpublished (1990)  
 Original source text: Hepatitis C virus, cDNA to viral RNA, clones K9-1 through 15e, isolated from chimpanzee (individual 910) blood plasma.  
 Draft entry and printed sequence for [1] kindly submitted by M. Houghton, 22-FEB-1990. Chiron Corporation, 4560 Horton Street, Emeryville CA 94608.  
 FEATURES  
 Location/Qualifiers  
 1..7310  
 /organism="Hepatitis C virus"  
 /mol\_type="genomic RNA"  
 /db\_xref="taxon:11103"  
 <1..57310  
 /note="polyprotein"  
 /codon\_start=3  
 /protein\_id="AA45677.1"  
 /db\_xref="GI:329875"  
 /translation="GCPERLASRPLTDPQGWGPISTYANGSPDQRPYCWHPKPKFC  
 GIVPAKSCGPFVCTFTSPVVTGRTDRSGAPTSWGENDDVFLNNRPLGNWFGC  
 TWNSTGFTKVCAGPFCVIGAGNNTLHCPTDCFRKHDPATYSCGSGPSTPLRCLVD  
 YPRLWHYPTCTINYTKIRMYGVGHEHLEAACNWRGERCDLERSELSPLLT  
 TQWQVLPCTFTLPALSTGLIHQNIVDVLYGVGSSIASWAIKWEYVLLFLLL  
 ADARVCSCLMMLLIQAERALENLIANAASLAGTGLVSLFVFCFANVKGWKP  
 GAVTYGHWPLLLLLLALPORAYALDTEVAASCGLVGLMALILSPYKRIKSWC  
 LQASLLKVPYFVRVQGLLFCALARKMIGHYVQVMIKLGALTGTGVYNNHLPRLD  
 WAHNGRDLAAVEPVFVSQMETKLTWAGDAAACGDIINGLPSVARRGRIILGPAD  
 GWYKGRLLAPITAYAOOTRGLGCIITSLTRDKNOVEGEVQIVSTAQTFLATCI  
 NGVCTVYHAGTRTIIASPKGVIOYNTVDQLVGWPAPQSGRSLSLTPCTCGSDLL  
 VALSTGTETPFYKGAIPLEVIKGRHLIFCHSKKCDLAALVALGINAVAYRGLD  
 KTHADVIPIRRRGRSGLSPISYLGKSGGFLPCPAGHAGVIFRAVCTRGVA  
 KAVDFIPVENLETTMRSPVFTDNSPPVFPQSFVAHLHAPTSGSGSKTKVPAAYAQG  
 YKVLNPSVAATLFGAYMSKAHGIDPMIRTVGRTITGSPITSTYSGFLADGGGS  
 GGAYDIIICDECHSDATSLIGTGLDQAEATAGARLVLATATPGSVTPVHPNIEE  
 VALSTGTETPFYKGAIPLEVIKGRHLIFCHSKKCDLAALVALGINAVAYRGLD  
 VSVIPTSGDVAVDALMTGVTGDFSDVDCNTQVTDVFSIDPTFTETLTPDQ  
 AVSRTORRGTGRGKGIYRFVAPGERPSPGMSVLCEDYDAGCAWELTPTVTR  
 LRAYMTWPGVQDHLFEWGVFTGLTHIDAHFLSQTKQSGENLPYLVAQATVCR  
 AQAPPSQPMWKKLLRIKPTLHPTGLYRLGAVQNEITLTHPTKYMTCMSADLE  
 VHTLWLVGVGLAALAAVCLSTGCVIVGVVLSGKPAIIPDREVLVYRFDMEBES  
 QWLPYTEQGMLEAQFKOKALGLLOTASRAQAEVIAVAVTNWOKLETFWAKHWNFTS  
 GTQYLAGLSTLPGNPAIASLMAFTAVTSLTSTLLENILGWAQAAQAAGAAATA  
 FVGAAGLAGAGSVGLKGLIDILAGTAGAGVAGALFAFKIMSGSEVSTEDVLLNLLPAI  
 LSPGALVGVCAAILRHVGPGEVAVQWNNRLIAFASRGNHVSPTHYPESDAARV  
 TALISLUTQLRLRHLQWISSECTTCCSGWLRIDWDITCEVLSDFKWLAKLMGTF  
 LPIGPSVORGKYGVRVDGIMHTRTCHGCAEITGHVNGTMRIGVPTCRNWMSTQF  
 PNYATVTPCTPLPAPNYTFAWRVSAEEVEIROYDQVPHVTGMTDNLKPCQVPS  
 PEFTLGDVRLHRPAPCKPLLRBSVFRVLGHEYPVGSQPLCEPPEVAVLISMLT  
 DPSHITAEAGRLARGSPSSASQSLKATCATCTANHDSKPEVDAEILANLLWR  
 YNPILVETWKKPDYEPVHVHGCPLPPKPPPPVPPVPPKRTVLTSTLSTALALATP  
 QEMGNITVETSENKVVILDSFPLVAEEDEREISVPAETILKRSRFAQALPWAPRD  
 SFSSSTGTIGDNTTTSSEPAKCPDSDAESYSSMPLEGEPPGDPDLDSLSWSVAC  
 SEANAEADVCCSMYSYTCALVTPCAAEQKLPINALSNLLRHNLVSTTSRSAC  
 QROKKTVPRLQVLDVSHYQDVLKEVAAASKVXANLLSVEEACSLIPPHSAKSKFGVG  
 AKDVRCHARKAVTHINSWKDLEDNVTPIIDTTIMAKNEVFCVQPEKGRKPARLIVF  
 PDLGVRVCKEMALYDVVTTLPLAVMSGSSYGFQSPQGRVEFLVQAWKSKKTPMGFSYD

TRCPDSTVTSIRTEBAIYQCCDLDPQARVAIKSLTERLYVGGPLNRSNGENCGYRR  
CRASGVLTSCGNLTCTCIKARAACRAAGLQDCTMLVCGDDLVIVICSAAGVQBDASL  
RAFTAMTRYASAPRDPQPEYDLELITSCSNVSAHDDGAGKRVYLTRDPPTPLAR  
AAWETAHTTPUNSHLNIIMFAPILMARMILMTHFVSILARDQLEQAUCREIYGACY  
SIEPLDPLPIIQRLL"

## ORIGIN

Query Match	86.7%;	Score 1786;	DB 14;	Length 7310;
Best Local Similarity	91.7%;	Pred. No. 0;		
Matches 1888;	Conservative	0;	Mismatches 170;	Indels 0; Gaps 0;

  

Qy	2	TGGGCGCTATACAGCGCTATGCCCAGCAGACAAGGGGCGCTTTGGGATGCAATACCA 61
Db	1729	TGGGCGCCATCACGCGCTAGCCCCAGCAGACAAGGGGCGCTTCTAGGGTGCAATACCA 1788
Qy	62	GCTTGACCGCGCGGACAAAACAGGTGGAGGGTGAGTTCCAGATCGTGTCAACTGCTG 121
Db	1789	GCCTAACTGGCGCGGACAAAACAGGTGGAGGGTGAGTTCCAGATCGTGTCAACTGCTG 1848
Qy	122	CCAGACTTTCTTGGCAACCTGCATTAACGGGGTGCTGGACTGTCTACCATGGAGCGG 181
Db	1849	CCAAAACCTTCTTGGCAACCTGCATTAACGGGGTGCTGGACTGTCTACCATGGAGCGG 1908
Qy	182	GAACAAGGACCATTTGGCTGCTAAGGGTCTGTATCCAGATGTACAACAATGTGAGCC 241
Db	1909	GAACGAGGACCATCGCTGCTACCAAGGGTCTGTATCCAGATGTATACCAATGTAGACC 1968
Qy	242	AAGACCTGTAAGGTGCGCGCTCCCAAGGTGCGCGCTCAITTAACACCATGCACTTGG 301
Db	1969	AAGACCTTGTGGGTGCGCGCTCCCAAGGTGCGCGCTCAITTAACACCATGCACTTGG 2028
Qy	302	GCTCTCGGACCTTTACTCTGTGTCACAGGACGCGCGATGTCATTCCTGTGCGCGGACGG 361
Db	2029	GCTCTCGGACCTTTACTCTGTGTCACAGGACGCGCGATGTCATTCCTGTGCGCGGAG 2088
Qy	362	GTATGACAGGGGAGCGCTGCTTTGGCCCGCGCTATCTCTTAATGAAAGGTCTCTCGG 421
Db	2089	GTATAGACGGGGAGCGCTGCTGTGCGCCCGCGCTATTCCTACTTGAAGGCTCTCTCG 2148
Qy	422	GAGGCGCTCTGTGTGCGCGCAGGACATGCGGTAGGCAATTCAGAGCGCGGATGCA 481
Db	2149	GGGGTCCGCTGTGTGCGCGCGGAGCAACCGGTGGGCATATTTAGGGCGCGGTGTGA 2208
Qy	482	CCCGTGGAGTGGTAAAGCGGCTGGACTTCAATCCCGTAGAGACTTAGAGACAACCATGA 541
Db	2209	CCCGTGGAGTGGTAAAGCGGCTGGACTTTATCCCTGTGGAGAACCTTAGAGACAACCATGA 2268
Qy	542	GGTCCCGGTGTCTCAGACAATCTCTCCCAACAGAGTGCCCCAGAGCTACCAAGTGG 601
Db	2269	GGTCCCGGTGTCTCAGGATAACTCTCTCCACAGTAGTGCCCCAGAGCTTCCAGGTGG 2328
Qy	602	CCCACTGTGATGTCCCAACCGGAGCGGTAAAGACCAAGGTCCCGCGCGCATACGACG 661
Db	2329	CTCACTCCATGTCTCCCAACGGAGCGGCAAGACCAAGGTCCCGCGCTGCATATGACG 2388
Qy	662	CTCAGGCTCAAGAGTGTGTGCTCAACCCCTCGTGTGCTGCAACAATGGGCTTTGGTG 721
Db	2389	CTCAGGCTCAAGAGTGTGTGCTCAACCCCTCTGTGTGCTGCAACAATGGGCTTTGGTG 2448
Qy	722	CTTACATGTCCAGGCGCCATGGATTTGATCTTAATCATCAGGACTGGGTGAGGACAATTA 781
Db	2449	CTTACATGTCCAGGCTCAAGGATCGATCTTAACATCAGGACCGGGGTGAGACAATTA 2508
Qy	782	CTACTGGCAGCCCGATCACGTATTTCACTACGGCAAGTTCTTTCGCGAGCGGGGTGT 841
Db	2509	CCACTGGCAGCCCGATCACGTATTTCACTACGGCAAGTTCTTTCGCGAGCGGGGTGT 2568
Qy	842	CAGGGGTGCTTATGACAATAATTTGTGACGAGTGCCACTCCACGGATGCAACATCCA 901
Db	2569	CAGGGGTGCTTATGACAATAATTTGTGACGAGTGCCACTCCACGGATGCCACATCCA 2628
Qy	902	TCTTGGGCATTGGCACTGTCTTGAACAAGCAGAGACCGGGGGGCGAGCTGACTGTGCG 961
Db		

Db	2629	TCCTTGGGCATCGSCACTGTCTCTTGAACAAGCAGAGACTCGGGGGCGAGACTGCTGTTGTC 2688
Qy	962	TGCCACCGGTACCCCTCGGGGCTCGTCACTGTGCCCCATCTTAACATCGAGAGGTTG 1021
Db	2689	TGCCACCGGCACCCCTCGGGGCTCGTCACTGTGCCCCATCTTAACATCGAGAGGTTG 2748
Qy	1022	CTCTGTCCCACTACCGGAGAGATCCCTTTTATGGAAGGCTATTCCTCTTGAAGCAATTA 1081
Db	2749	CTCTGTCCCACTACCGGAGAGATCCCTTTTATGGAAGGCTATTCCTCTTGAAGCAATTA 2808
Qy	1082	AGGGGGGAGACATCTCATCTTCTGCCACTCAAAGAAAGTGCAGCAGCTCGCCGCAA 1141
Db	2809	AGGGGGGAGACATCTCATCTTCTGCCACTCAAAGAAAGTGCAGCAGCTCGCCGCAA 2868
Qy	1142	NACTGTCGCGTGGCGTCAATGCGGTCTTACTACCGCGGCTTGTGATGTCTCGTCA 1201
Db	2869	AGCTGTGTCATTTGGGCATCAATGCGGTGCGCTTACTACCGCGGCTTGTGATGTCTCGTCA 2928
Qy	1202	TCCGACCACTGTGACGTTGTCTGTGGCAACTGACGCCCTCATGACCGGCTTTACCG 1261
Db	2929	TCCGACCACTGTGACGTTGTCTGTGGCAACTGACGCCCTCATGACCGGCTTTACCG 2988
Qy	1262	GGACTTTCGATTCGCTGATAGACTGCAACACGTTGTCTACCCAGACAGTCTCAGCC 1321
Db	2989	GGACTTTCGACTCGGTGATAGACTGCAATACGTTGTCTACCCAGACAGTCTCAGCC 3048
Qy	1322	TTGACCCCTACCTTCACTTGGAGACATCAACGTTTCCCGAGGATGCTGTCTCCGCTACTC 1381
Db	3049	TTGACCCCTACCTTCACTTGGAGACATCAACGTTTCCCGAGGATGCTGTCTCCGCTACTC 3108
Qy	1382	AACGTCGGGGTAGGACTGGCAGAGGAAAGCAGGATCTACAGATTTGTGGCAACCGGGG 1441
Db	3109	AACGTCGGGGTAGGACTGGCAGAGGAAAGCAGGATCTACAGATTTGTGGCAACCGGGG 3168
Qy	1442	AGGTCCTTCTTGGCATGTTTGAATCTGTCTCTGTGAGTGTATGACGCGGTTGTG 1501
Db	3169	AGGTCCTTCTTGGCATGTTTGAATCTGTCTCTGTGAGTGTATGACGCGGTTGTG 3228
Qy	1502	CTTGTGTAGACTTACGCCCGCGAGACACAGTTAGGCTACAGCATATGACACACC 1561
Db	3229	CTTGTGTAGACTTACGCCCGCGAGACACAGTTAGGCTACAGCATATGACACACC 3288
Qy	1562	CGGGACTTCCCGTGTGCCAAGACCATCTTTGAATTTTGGAGGCGCTCTTTACGGGTCTCA 1621
Db	3289	CGGGACTTCCCGTGTGCCAAGACCATCTTTGAATTTTGGAGGCGCTCTTTACGGGTCTCA 3348
Qy	1622	CCCATAGACGCCCATCTTCTATCCAGACAAAGAGGTGGGAAACCTTCCCTATC 1681
Db	3349	CTCATATAGATCCCATCTTCTATCCAGACAAAGAGGTGGGAAACCTTCCCTATC 3408
Qy	1682	TGCTAGCTTACCAAGCCCGTGTGCTTAGAGCTCAAGCCCTCCCGTCTCGTGGGACC 1741
Db	3409	TGCTAGCTTACCAAGCCCGTGTGCTTAGAGCTCAAGCCCTCCCGTCTCGTGGGACC 3468
Qy	1742	AGATGTGAAGTGTCTGATTCGCTCTCAAGCCCAACCTCCATGGGCAACACCTCTGCTAT 1801
Db	3469	AGATGTGAAGTGTCTGATTCGCTCTCAAGCCCAACCTCCATGGGCAACACCTCTGCTAT 3528
Qy	1802	ATAGACTGGGCGCTCTCAGAAATGAAGTCAACCTGACGCAACCTGACCAAGTATATCA 1861
Db	3529	ACAGACTGGGCGCTCTCAGAAATGAAGTCAACCTGACGCAACCTGACCAAGTATATCA 3588
Qy	1862	TGACATGTATGTGCGCTGACCTGGAGGTCTGTCACAGTACCTGGGTGCTCGTTGGCGGG 1921
Db	3589	TGACATGTATGTGCGCGAGCTGGAGGTCTGTCAGAGCACTGGGTGCTCGTTGGCGGG 3648
Qy	1922	TTCTGGCTCTTGGCCCGCTATTGCTTATCCAGGCTGCTGCTCATAGTAGGATGAGA 1981
Db	3649	TCTGGCTCTTGGCCCGCTATTGCTTATCCAGGCTGCTGCTCATAGTAGGATGAGA 3708
Qy	1982	TTGTCTTGTCCGGAAGCGGCAATCATACCGACAGAGAGTCTCTTACCGGAGTTGCG 2041
Db	3709	TGCTCTTGTCCGGGAAGCGGCAATCATACCTGACAGGGAAGTCTCTTACCGAGAGTTGCG 3768



Qy 2042 ATGAAATGGAAGAGTGCT 2059  
| | | | |  
Db 3769 ATGAGATGGAAGAGTGCT 3786

RESULT 13  
108294  
LOCUS 108294 9185 bp DNA linear PAT 02-DEC-1994  
DEFINITION Sequence 1 from Patent EP 0388232.  
ACCESSION 108294  
VERSION 108294.1 GI:588994  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 9185)  
AUTHORS Houghton, M., Choo, Q.-L. and Kuo, G.  
TITLE NABV diagnostics and vaccines  
JOURNAL Patent: EP 0388232-A1 1 19-SEP-1990;  
FEATURES  
Location/Qualifiers  
source 1. 9185  
/organism="unknown"  
/mol\_type="unassigned DNA"

ORIGIN  
Query Match 86.7%; Score 1786; DB 6; Length 9185;  
Best Local Similarity 91.7%; Pred. No. 0;  
Matches 1888; Conservative 0; Mismatches 170; Indels 0; Gaps 0;

Qy 2 TGGGCGCTATCACGGCCTATGCCAGAGATGCCAGAGACAAGGGGCTTTGGGATGATATACCA 61  
| | | | |  
Db 3396 TGGGCGCCATCACGGCGTACGCCAGCAGACAAGGGGCTCTAGGGTGCATATACCA 3455

Qy 62 GCTTGACCGCGCGGACAAACACAGGTGGAGGTGAGTTTCAGATCGTCAACTGCTG 121  
| | | | |  
Db 3456 GCCTAACTGGCGCGGACAAACACAGTGGAGGTGAGGTCCAGATTGTGCAACTGCTG 3515

Qy 122 CCCAGACTTTCTTGGCAACTGCATTAACGGGGTGTGTGGACTGTCTACCATGGAGCCG 181  
| | | | |  
Db 3516 CCCAAACCTTCTGGCAACGTGCATCAATGGGGTGTGTGGACTGTCTACCAAGGGGCG 3575

Qy 182 GAACAAGACCATTTGGCTACCTAAGGCTCTGTTATCCAGATGATACCAATGTGAACC 241  
| | | | |  
Db 3576 GAACAGGACCATCGCGTACCCAGGGTCTCTGCATCCAGATGATATACCAATGTAGACC 3635

Qy 242 AAGACCTCGTAGGCTGCCCGCTCCCAAGGTGCCCTCATTAACACCATGCACTTGGC 301  
| | | | |  
Db 3636 AAGACCTTGTGGGTGGCGCGCTCCGCAAGGTAGCCGCTCAATTGACACCCCTGCACTTGGC 3695

Qy 302 GCTCCTCGGACCTTTACTGTGTACAGGACGCGCGATGTCATTCTCTGTGCGCGACGCGG 361  
| | | | |  
Db 3696 GCTCCTCGGACCTTTACTGTGTACAGGACGCGCGATGTCATTCCGTGCGCGCGCGG 3755

Qy 362 GTGATGACGGGCGAGCTGTCTTGGCCCGGCTATCTCTTACTTGAAGGCTCTCTCGG 421  
| | | | |  
Db 3756 GTGATAGCAGGGGCGAGCTGTGTGCGCCCGGCCCATTTCTACTTTGAAGGCTCTCTCGG 3815

Qy 422 GAGCCCTCTGCTGTGCCCGCAGACATGCCGTAGGCATATTCAGAGCCCGGTATGCA 481  
| | | | |  
Db 3816 GGGGTCCGCTGTGTGCCCGCGGGCAAGCCGTGGGCATATTTAGGGCGCGGTGTGCA 3875

Qy 482 CCCGTGAGTGGCTAAGCGGTGACTTTCAATCCCGGTAGAGAGCTTAGAGACCAACATGA 541  
| | | | |  
Db 3876 CCCGTGAGTGGCTAAGCGGTGACTTTATCTCTGTGGAGAACCTAGAGACCAACATGA 3935

Qy 542 GGTCCCCTGTTCTTACAGACAATCTCTCCCAACAGCAGTGCCTCCAGAGCTTACCAAGTGG 601  
| | | | |  
Db 3936 GGTCCCCTGTTTCAACGATAACTCTCTCCACAGTAGTGCCTCCAGAGCTTCCAGGTGG 3995

Qy 602 CCACCTGCATGCTCCCAACGGGAGCGGTAGAGCAACAAAGTCCCGGCGCATACGAG 661  
| | | | |  
Db 3996 CTCACCTCCATGCTCCCAACGGGAGCGGCAAGACCAAGGTCCCGGCTGCATATGAG 4055

Qy 662 CTCAGGGCTACAAAGTGCTGCTCAACCCCTCCGTTGCTGCAACAATGGCTTTGGTG 721  
| | | | |  
Db 4056 CTCAGGGCTATAAGGTGCTAGTACTCAACCCCTCTGTTGCTGCAACAATGGCTTTGGTG 4115

Qy 722 CTTACATGTCCAAGGCCCATGGGATTCATCTTAACATCAGGACTGGGGTGAGGACAATTA 781  
| | | | |  
Db 4116 CTTACATGTCCAAGGGCTCATGGGATCGATCTTAACATCAGGACCGGGGTGAGAACAAATTA 4175

Qy 782 CTACTGCGACCCGATCACGTATTTCACCTACGCAAGTTCCTTGGCCGACGGCGGTGTT 841  
| | | | |  
Db 4176 CCACTGGCAGCCCATCACGTACTTCACCTACGCAAGTTCCTTGGCCGACGGCGGTGCT 4235

Qy 842 CAGGGGTGCTTATGACATATAATTTGTGACGAGTCCCACTCCACGGATGCACATCCA 901  
| | | | |  
Db 4236 CGGGGGCGCTTATGACATATAATTTGTGACGAGTGCCTCCACGGATGCCACATCCA 4295

Qy 902 TCTTGGCATTTGGCATTTCTTTGACCAAGCAGAGACCGGGGGCGAGACTGACTGTGC 961  
| | | | |  
Db 4296 TCTTGGCATTTGGCATTTCTTTGACCAAGCAGAGACTGGGGGGCGAGACTGCTTTGTGC 4355

Qy 962 TCGCACCCGCTACCCCTCCGGGCTCCGTCACGTGCCCCATCTTAACATCGAGGAGTTG 1021  
| | | | |  
Db 4356 TCGCACCCGCTACCCCTCCGGGCTCCGTCACGTGCCCCATCTTAACATCGAGGAGTTG 4415

Qy 1022 CTCTGTCACTACCGGAGAGATCCCTTTTATGGCAAGGCTATTCCCTTTGAAGCAATTA 1081  
| | | | |  
Db 4416 CTCTGTCACTACCGGAGAGATCCCTTTTATGGCAAGGCTATTCCCTTTGAAGTAATCA 4475

Qy 1082 AGGGGGGAGACATCTCATCTTCTGCCACTCAAAGAAGAGTGGGAGAGCTGCCCGCAA 1141  
| | | | |  
Db 4476 AGGGGGGAGACATCTCATCTTCTGTCAATCAAAGAAGAGTGGGAGAGCTGCCCGCAA 4535

Qy 1142 AACTGTGCGGTTGGGCGTCAATGCGTGCTTACTACCGCGCTTGTGATGTGCGTCA 1201  
| | | | |  
Db 4536 AGTGTGCGATTTGGGCATCAATGCCGTGCCCTACTACCGCGTCTTGAAGTGTCCGTCA 4595

Qy 1202 TCCGACCACTGTGTGCTGCTGTGGCAACTGACGCCCTCATGACCGGCTTTTACCG 1261  
| | | | |  
Db 4596 TCCGACCACTGTGTGCTGCTGTGGCAACCGATGCCCTCATGACCGGCTTATACCG 4655

Qy 1262 GCACCTTCGATTCGGTGTATGACTGMAACAGTGTGTGACCCAGACAGTGCATTCAGCC 1321  
| | | | |  
Db 4656 GCACCTTCGATTCGGTGTATGACTGCAATAGTGTGTACCCAGACAGTGCATTCAGCC 4715

Qy 1322 TTGACCCCTACTTCCACATTTGACACATCAAGTTCGCCAGGATGCTGTCCCGTACTC 1381  
| | | | |  
Db 4716 TTGACCCCTACTTCCACATTTGACACATCAAGTTCGCCAGGATGCTGTCTCCCGACTC 4775

Qy 1382 AACGTGGGGTAGGACTGGCAGAGGGAAGCCAGGCATCTACAGATTTGTGGCACCGGGGG 1441  
| | | | |  
Db 4776 AACGTGGGGTAGGACTGGCAGAGGGAAGCCAGGCATCTACAGATTTGTGGCACCGGGGG 4835

Qy 1442 AGCGTCTTCTGGGATGTTTGAAGTGTGTCTCTGCGAGTGTGATGACGCGGGTGTG 1501  
| | | | |  
Db 4836 AGCGTCTTCTGGGATGTTTGAAGTGTGTCTCTGCGAGTGTGATGACGCGGGTGTG 4895

Qy 1502 CTTTGTATGAGCTTACGCCCGCAGACACAGTTAGGCTACGAGCATACATGAACACC 1561  
| | | | |  
Db 4896 CTTTGTATGAGCTTACGCCCGCAGACACAGTTAGGCTACGAGCGGTATGATGAACACC 4955

Qy 1562 CGGGACTTCCCGTGTGCCAAGACCATCTTGAAATTTTGGGAGGGCGCTTTTACGGGTCTCA 1621  
| | | | |  
Db 4956 CGGGACTTCCCGTGTGCCAAGACCATCTTGAAATTTTGGGAGGGCGCTTTTACGGGTCTCA 5015

Qy 1622,CCCAATAGACGCCCATCTTCTATCCAGACAAAGCAGAGTGGGGAACCTTCCCTATC 1681  
| | | | |  
Db 5016 CTATATAGATGCCCATCTTCTATCCAGACAAAGCAGAGTGGGGAACCTTCCCTATC 5075

Qy 1682 TGGTAGGCTACCAAGCAGCGTGTGGCTAGAGCTCAAGCCCTCCCGCTCGTGGGACC 1741  
| | | | |  
Db 5076 TGGTAGGCTACCAAGCAGCGTGTGGCTAGAGCTCAAGCCCTCCCGCTCGTGGGACC 5135

QY 1742 AGATGTGAAGTGTCTGATCGTCTCAAGCCACCCCTCCATGGGCCAACACCTCTGCTAT 1801  
Db 5136 AGATGTGAAGTGTCTGATTCGCTCAAGCCACCCCTCCATGGGCCAACACCCCTGCTAT 5195  
QY 1802 ATAGACTGGGCGCTGTCCAGAAATGAAGTACCCCTGACGACCCAGTACCAAGTATATCA 1861  
Db 5196 ACAGACTGGGCGCTGTTCAGAAATGAAGTACCCCTGACGACCCAGTACCAAAATACATCA 5255  
QY 1862 TGACATGTATCTCGGCTGACCTGAGGTGCGTACAGAGTACCTGGTGTCTGTTGGCGCG 1921  
Db 5256 TGACATGTATCTCGGCGACCTGAGGTGCGTACAGAGTACCTGGTGTCTGTTGGCGCG 5315  
QY 1922 TTCTGGCTGTCTTGGCGCGCTATTGCGCTATCCACAGGCTGCGTGTCTATAGTAGTAGGA 1981  
Db 5316 TCCTGGCTGTCTTGGCGCGCTATTGCGCTGTCAACAGGCTGCGTGTCTATAGTAGTAGGA 5375  
QY 1982 TTGTCTTGTCCGAAAGCGCGCAATCATACCCGACGGAAGTCTCTACCGGAGTTG 2041  
Db 5376 TCGTCTTGTCCGGAAGCGCGCAATCATACCTGACGGAAGTCTCTACCGAGAGTTG 5435  
QY 2042 ATGAATGGAAGAGTGTCT 2059  
Db 5436 ATGAGATGGAAGAGTGTCT 5453

RESULT 14  
BD091382 9185 bp DNA linear PAT 27-AUG-2002  
LOCUS  
DEFINITION HCV cultivation method in eucaryotic cells.  
ACCESSION BD091382  
VERSION BD091382.1 GI:22636993  
KEYWORDS JP 2001314192-A/3.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 9185)  
AUTHORS Weiner,A.J., Steimer,K.S. and Houghton,M.  
TITLE HCV cultivation method in eucaryotic cells  
JOURNAL Patent: JP 2001314192-A 3 13-NOV-2001;  
CHIRON CORP  
OS Homo sapiens (human)  
PN JP 2001314192-A/3  
PD 13-NOV-2001  
PF 15-MAR-2001 JP 2001075114  
PR 25-AUG-1989 US 398667  
PI AMY J WEINER,KATHELYN S STEIMER,MICHAEL HOUGHTON PC  
C12N15/09,C12N5/10,C12N7/00//C12N15/00,C12N15/00, C12N5/00  
CC HCV cultivation method in eucaryotic cells  
FH Key Location/Qualifiers  
FT source 1..9185  
FT Location/Qualifiers  
1..9185 /organism='Homo sapiens (human)'.  
/organism='Homo sapiens'  
/mol\_type='genomic DNA'  
/db\_xref='taxon:9606'

FEATURES  
source  
1..9185  
Location/Qualifiers  
1..9185 /organism='Homo sapiens (human)'.  
/organism='Homo sapiens'  
/mol\_type='genomic DNA'  
/db\_xref='taxon:9606'

ORIGIN  
Query Match 86.7%; Score 1786; DB 6; Length 9185;  
Best Local Similarity 91.7%; Pred. No. 0;  
Matches 1888; Conservative 0; Mismatches 170; Indels 0; Gaps 0;  
QY 2 TGGCGCTATCAGCGCTATGCCAGCAGCAAGAGGCGCTTTGGGATGCAATACCA 61  
Db 3396 TGGCGCGCATCAGCGCGTACGCCAGCAGCAAGAGGCGCTCTTAGGGTGTCAATACCA 3455  
QY 62 GCTTGACCGCGCGGCAAAACACAGTGGAGGCTGAGGTTCAAGTCTCACTGCTG 121  
Db 3456 GCCTACTGCGCGGCAAAACACAGTGGAGGCTGAGGTTCAAGTCTCACTGCTG 121  
QY 122 CCAGACTTTCTTGGCAACCTGCAATTAACGGGGTGTGTTGGACTGTCTACCATGGAGCG 181

Db 3516 CCACAACTTCTCTGGCAAAGTGCATCAATGGGGTGTCTGGACTGTCTACCAAGGGCGG 3575  
QY 182 GAACAAGGACCAATGTGCGTCACTTAAGGTCCTGTTATCAGATGTATACCAATGTGAC 241  
Db 3576 GAACGAGGACCAATGTGCGTCACTTAAGGTCCTGTTATCAGATGTATACCAATGTGAC 3635  
QY 242 AAGACCTCGTAGGTCGGCGCTCCCAAGGTGCGCGCTCATTAACACCATGCACTTGG 301  
Db 3636 AAGACCTTGTGGGCTGGCGCGCTCCCAAGGTGAGCGCGCTCATTAACACCATGCACTTGG 3695  
QY 302 GCTCTCTCGACCTTTACTGTGTACGAGGACGCGCGATGTCAATTCGTGTGCGCGG 361  
Db 3696 GCTCTCTCGACCTTTACTGTGTACGAGGACGCGCGATGTCAATTCGTGTGCGCGG 3755  
QY 362 GTGATGGCAGGGGACGCTGTCTTTGCCCGCGCTATCTCTTACTTGAAGGCTCTCTCG 421  
Db 3756 GTGATAGCAGGGGACGCTGTCTTTGCCCGCGCTATCTCTTACTTGAAGGCTCTCTCG 3815  
QY 422 GAGGCGCTCTGCTGTGCGCGGACGAGCATGCGTAGGCAATATTCAGAGCGCGGTATGCA 481  
Db 3816 GGGGTCCGCTGTGTGCGCGCGGACGCGCTGGGCAATATTAAGGCGCGGTGTGCA 3875  
QY 482 CCGGTGGAGTGTGAAGGCGGTGGAATTCATCCCGTAGAGACTTTAGAGACCAACCATCA 541  
Db 3876 CCGGTGGAGTGTGAAGGCGGTGGAATTCATCCCGTAGAGAACTTAGAGACCAACCATCA 3935  
QY 542 GGTCCCGGTGTCTCAGACAACTCTCTCCACAGCAGTGTCCCGCAGAGTACCAAGTGG 601  
Db 3936 GGTCCCGGTGTCTCAGGATACTCTCTCCACAGTAGTGTCCCGCAGAGTCTCCAGTGG 3995  
QY 602 CCCACTGTCATCTCCACCGCGGAGGTAAAGACGACCAAGGTCCCGCGCGCATAGCGAG 661  
Db 3996 CTCACCTCATCTCCACAGGCGGCAAGACCAAGGTCCCGCGGTGCATATGCA 4055  
QY 662 CTCAGGGCTACAGGTGTGTGTCTCAACCCCTCGTGTCTCAACAAATGGGCTTTGGT 721  
Db 4056 CTCAGGGCTATAGGTGTGTGTCTCAACCCCTCTGTGTCTCAACACTGGGCTTTGGT 4115  
QY 722 CTTATCATGTCCAAAGGCCAATGGGATGTATCTTAAATCAGAGTGTGGGTGAGGACAAATTA 781  
Db 4116 CTTATCATGTCCAAAGGTGTATGGGATGTATCTTAAATCAGAGTGTGGGTGAGGACAAATTA 4175  
QY 782 CTACTGGCAGCGCGATCAGTATTCACCTACGCGCAAGTCTCTCCGCGCGCGGTGT 841  
Db 4176 CCAGTGGCAGCGCGCATCAGTATTCACCTACGCGCAAGTCTCTCCGCGCGCGGTGT 4235  
QY 842 CAGGGGGTCTTATGACATAATAATTTGTGACGAGTGTCCACGAGTGTCCACGAGTGTCCCA 901  
Db 4236 CAGGGGGTCTTATGACATAATAATTTGTGACGAGTGTCCACGAGTGTCCCA 4295  
QY 902 TCTTGGGCAATGGCACTGTCTTGAACCAAGCAGAGACCGCGGGGCGGAGTGTGTCG 961  
Db 4296 TCTTGGGCAATGGCACTGTCTTGAACCAAGCAGAGTGTCCGCGGGGCGGAGTGTGTCG 4355  
QY 962 TCGCCACCGCTACCCCTCGGGCTCCGCTACGCTGTCGCCCATCTAACATCGAGGAGTGT 1021  
Db 4356 TCGCCACCGCTACCCCTCGGGCTCCGCTACGCTGTCGCCCATCTAACATCGAGGAGTGT 4415  
QY 1022 CTCTGTCCACTACCGGAGAGATCCCTTTTATGGCAAGGCTATTCCTCTTGAAGCAATTA 1081  
Db 4416 CTCTGTCCACTACCGGAGAGATCCCTTTTATGGCAAGGCTATTCCTCTTGAAGCAATTA 4475  
QY 1082 AGGGGGGAGACATCTCTCTTCTGCCACTAAAGAGAGTGTGCGAGCTGTGCGGCA 1141  
Db 4476 AGGGGGGAGACATCTCTCTTCTGCCACTAAAGAGAGTGTGCGAGCTGTGCGGCA 4535  
QY 1142 AACTGCTCGGTGGCGCTCAATGCGGGCTTACTACCGGGCTTCAATGTGTCCGCA 1201  
Db 4536 AGCTGCTCGATTTGGGCAATGATGCGGTCTACTACCGGTCTTCACTGTGTCCGCA 4595  
QY 1202 TCCCGACAGTGTGTGACGTTGTGTGTGCGGCACTGACGCGCTCATGACCGGCTTTACG 1261

4596	Db	TCCGACACAGCGGCGATGTTGCTGCTGGGCAACCGATGCCCTCATGACCGGCTATACCG	4655
1262	Qy	GCGACTTTCGATTTCGGTGATAGACTGCAACACGCTGTGTACACCCAGACAGTCGACTTCAGCC	1321
4656	Db	GCGACTTCGACTCGGTGATAGACTGCAATACGTGTGTACCCACACAGAGTCGATTTCAGCC	4715
1322	Qy	TTGACCTCTACCTTCACCATTTGAGACAATCAGCTTCCCCAGGATGCTGTCTCCCGTACTC	1381
4716	Db	TTGACCTCTACCTTCACCATTTGAGACAATCAGCTTCCCGCAGGATGCTGTCTCCCGCATC	4775
1382	Qy	AACGTCGGGGTAGGACTGGCAGAGGGGAAGCCAGGCATCTCAGAGATTTGTGGCACCGGGG	1441
4776	Db	AACGTCGGGGCAGACTGGCAGGGGGGAAGCCAGGCATCTCAGAGATTTGTGGCACCGGGG	4835
1442	Qy	AGGTCCTCTTTGGGATGTTTGACTCGTCTGTCTCTCTGCGAGTGCTATGACGCGGGTTGTG	1501
4836	Db	AGCCTCCCTCCGGCATGTTGACTCGTCTGTCTCTGTGAGTGTATGACGCAAGCTGTG	4895
1502	Qy	CTTGCTATGAGCTTACGCCGCCGAGACCAACAGTTAGGCTACGAGCATACATGAACACC	1561
4896	Db	CTTGCTATGAGCTACGCCGCCGAGACTACAGTTAGGCTACGAGCGTACATGAACACC	4955
1562	Qy	CGGCACTTCGCGTGTGCCAAGACCATCTTGAAATTTTGGGAGGCGCTCTTTACGGGTCTCA	1621
4956	Db	CGGGCTTCCGCTGTGCCAGSACCATCTTGAAATTTTGGGAGGCGCTCTTTACAGGCTCA	5015
1622	Qy	CCGACATAGACGCCACTTCTCTATCCAGACAAGAGAGTGGGGAAAACTTCCCTATC	1681
5016	Db	CTCATATAGATGCCCACTTCTATCCAGACAAGAGAGTGGGGGAAACCTTCTCTTACC	5075
1682	Qy	TGGTAGCGGTACCAAGCACCGTGTGCGCTAGAGCTCAAGCGCCCTCCCGCGTGGGAGCC	1741
5076	Db	TGGTAGCGGTACCAAGCACCGTGTGCGCTAGGCGTCAAGCGCCCTCCCGCATCGTGGAGCC	5135
1742	Qy	AGATGTGGAGTGTCTTGATTCGTTCTAAGGCCACCCCTCCATGGGCCAACCTCTGCTAT	1801
5136	Db	AGATGTGGAGTGTCTTGATTCGCTCTAAGGCCACCCCTCCATGGGCCAACCCCTGCTAT	5195
1802	Qy	ATAGACTGGGCGCTGCAGAACTGAAGTCAACCTTGACGCCACCGAGTCAACAGTATATCA	1861
5196	Db	ACAGACTGGGCGCTGTTCAGAAATGAATACCCCTGACGCCACCGAGTCAACAAATATATCA	5255
1862	Qy	TGACATGTATGTTCGGCTGACCTGGAGGTCTGTCAGAGTACCTGGGTGCTCGTTGGCGGCG	1921
5256	Db	TGACATGTATGTTCGGCGACCTGGAGGTCTGTCAGAGCACTGGGTGCTCGTTGGCGGCG	5315
1922	Qy	TTCTGGCTGCTTTGGCGCGGTATTGCTATTCACAGCGTGGTGGTCAATAGTAGTAGGA	1981
5316	Db	TCCTGGCTGCTTTGGCGCGGTATTGCTGTCAACAGGCTGGTGGTCAATAGTAGTAGGAGG	5375
1982	Qy	TTGTCTTGTTCGGAAAGCCGCAATCATACCCGACAGGGAAGTCTCTACCGGAGTTTCG	2041
5376	Db	TCGTCTTGTTCGGGAAGCCGCGCAATCATACCTGACAGGGAAGTCTCTACGAGAGTTTCG	5435
2042	Qy	ATGAAATGGAAGAGTGCT	2059
5436	Db	ATGAGATGGAAGAGTGCT	5453

RESULT 15					
AR166930					
LOCUS	AR166930	9379 bp	DNA	linear	PAT 17-OCT-2001
DEFINITION	Sequence 1 from patent US 6284249.				
ACCESSION	AR166930				
VERSION	AR166930.1	GI:16243325			
KEYWORDS	.				
SOURCE	Unknown.				
ORGANISM	Unknown.				
REFERENCE	Unclassified.				
AUTHORS	1 (bases 1 to 9379)				
TITLE	Barban, V.				
	Fusion polypeptide having the C protein and E1 protein of hepatitis C virus				

JOURNAL FEATURES	Patent: US 6284249-A 1 04-SEP-2001; Location/Qualifiers 1. .9379 /organism="unknown" /mol_type="unassigned DNA"	Query Match Best Local Similarity 86.7%; Score 1786; DB 6; Length 9379; Matches 1888; Conservative 0; Mismatches 170; Indels 0; Gaps 0;
Qy	2	TGGGGCCTATCACGGCCTATGCCAGAGACAGAGGGGCCTTTTGGGATGATATATCACCA 61
Db	3396	TGGCGCCCATACGGCGTACGCCAGCAGACAAGGGCCTCTAGGGTGATATATCACCA 3455
Qy	62	GCTTTCAGCGGCGGGACAAAACACAGGTGAGGGTGAGGTTTCAGATCGTGTCAACTGCTG 121
Db	3456	GCCTAACTGCGCGGGACAAAACCAAGTGGGGTGAGGTCAGATTTGTCAACTGCTG 3515
Qy	122	CCCAGACTTTCTTTGGCAACCTGCATTAAACGGGGTGTTTGGACTGTCTACCATGGAGCCG 181
Db	3516	CCCAAACTTTCTTGGCAACCTGCATCAATGGGGTGCTGACTGTCTACCCAGGGCCG 3575
Qy	182	GAACAAGGACCATTTGCGTCACTTAAGGGTCCTGTTATCCAGATGTACCAATGTGGAACC 241
Db	3576	GAACGAGGACCATCGCGCTCAACCAAGGGTCCTGTCTATCCAGATGTATACCAATGTAGACC 3635
Qy	242	AAGACCTCGTAGCTGGCCCGCTCCCCAAGGTGCCCGCTCATTTAAACCACTGACACTTGGC 301
Db	3636	AAGACCTTGTGGGTGGCCCGCTCCGCAAGGTAGCCGCTCATTTGACACCTGTGACTTGGC 3695
Qy	302	GCTCTCGGACCTTTACTGCTGACAGGCGACGCCGATGTCAATCTGTGCGCGCAGCGG 361
Db	3696	GCTCTCGGACCTTTACTTGGTCAOGAGGCAAGCCGATGTCAATCCGTGCGCGCGCGG 3755
Qy	362	GTGATGCAAGGGGAGCCTGTTTGGCCCGGCTATCTCTTACTTTGAAAGGCTCTCTCG 421
Db	3756	GTGATAGAGGGGAGCCTGCTGTGCGCCCGGCCCATTTCTACTTTGAAAGGCTCTCTCG 3815
Qy	422	GAGGCCCTCTGCTGTGCCCGCAGGACATGCGGTAGGCATATTCAGAGCCGCGGTATGCA 481
Db	3816	GGGGTCGCGCTGTGTGCCCCGCGGGGACGCCGTGGGCATATTTAGGGCCGCGGTGTGCA 3875
Qy	482	CCCGTGAGTGGCTTAAGGCGGTGACCTTCATCCCGGTAGAGAGCTTAGACACAACCATGA 541
Db	3876	CCCGTGAGTGGCTTAAGGCGGTGACCTTTATCCCTGTGGAGAACCTTAGACACAACCATGA 3935
Qy	542	GGTCCC CGGTGTTCTCAGACAACTCTCTCCCAACAGAGTGCCCCAGAGTACCAAGTGG 601
Db	3936	GGTCCC CGGTGTTCCAGGATAACTCTCTCTCCACAGTAGTGCCCAAGAGCTTCCAGGTG 3995
Qy	602	CCACCTGCATGCTCCACCGGAGCGGTAAAGACCAAGGTCCTCCGCGCGCATACGCAG 661
Db	3996	CTCACTTCCATGTCTCCACAGGAGCGGCAAAAGACCAAGGTCCTCCGCGTGCATATGCG 4055
Qy	662	CTCAGGCTACAAGGTGCTGGTGCTCAACCCCTCCGTTGCTGCAAACTGGGCTTTGGTG 721
Db	4056	CTCAGGCTATTAAGGTGCTAGTACTCAACCCCTCTGTTGCTGCAACTGGGCTTTGGTG 4115
Qy	722	CTTACATGTCGAAGGCCCATGGGATTGATCCTTAACATCAGGACTGGGGTGAGGACAATTA 781
Db	4116	CTTACATGTCGAAGGCTCATGGGATCGATCCTTAACATCAGGACCGGGGTGAGAACAAATTA 4175
Qy	782	CTACTGGCAGCCGATCATCGTATTTCCACCTACGGCAAGTTCCTTGGCGAGCGGGGTGTT 841
Db	4176	CCACTGGCAGCCCGCATCACGCTACTCCACCTACGGCAAGTTCCTTGGCGAGCGGGGTGCT 4235
Qy	842	CAGGGGTGCTTATGACATAATAATTTGTGACAGGTGCCACTCCACGGATGCAACATCCA 901
Db	4236	CAGGGGTGCTTATGACATAATAATTTGTGACAGGTGCCACTCCACGGATGCCACATCCA 4295
Qy	902	TCTTGGGCATTGGCACTGTCTTTCACCAAGCAGAGACCGCGGGGGGAGACTGACTGTGC 961

Db	4296	TCCTGGGACATGGGACATGTCCTTGACCAAGCAGAGACTGCGGGGCGAGACTGTTGTGC	4355
Qy	962	TCGCGACCGCTACCCCTCGGGCTCCGTCACCTGTCGCCCATCTTAACATCGAGAGGTTG	1021
Db	4356	TCGCGACCGCACCCCTCCGGGCTCCGTCACCTGTCGCCCATCTTAACATCGAGAGGTTG	4415
Qy	1022	CTCTGTCCACTACCGGAGAGATCCCTTTTATGGAAGGCTATTCCTTTGAAGCAATTA	1081
Db	4416	CTCTGTCCACTACCGGAGAGATCCCTTTTACGSCAAGGCTATCCCTCGAAGTAATCA	4475
Qy	1082	AGGGGGGAGACATCTCTCATCTTCTGCCACTCAAGAAGAGTGCAGAGCTCGCGCAA	1141
Db	4476	AGGGGGGAGACATCTCTCATCTTCTGTCAATTCAGAGAGAGTGCAGAACTCGCGCAA	4535
Qy	1142	AACGTGTCGGCTTGGGGCTCAATGCGCTGGCTTACTACGCGGCTTGATGTGTCGTCA	1201
Db	4536	AGCTGTGCGATTGGGGCATCAATGCCGTGGCTACTACCGGGCTTGGAGCTGTCGTCA	4595
Qy	1202	TCCGACCAAGTGTGACGTTGTGTCGTGGCACTGACGCGCTCATGACCGGCTTTACCG	1261
Db	4596	TCCGACCAAGCGGCGATGTTGTGTCGTGGCAACCGATGCCCTCATGACCGGCTATACCG	4655
Qy	1262	GCGACTTCGATTTCGGTGATAGACTGCAACAGGTGTGTACCCAGACAGTTCAGCTT	1321
Db	4656	GCGACTTCGACTTCGGTGATAGACTGCAATACGTGTGTACCCAGACAGTTCAGCTT	4715
Qy	1322	TTGACCCCTACCTTCACCATTTGAGACAATCACGCTTCCCGAGGATGCTGTCTCCCGTACTC	1381
Db	4716	TTGACCCCTACCTTCACCATTTGAGACAATCACGCTTCCCGAGGATGCTGTCTCCCGACTC	4775
Qy	1382	AACGTTCGGGTAGGACTGGCAGAGGGAAGCCAGGCAATCTACAGATTTGTGGCACCGGGG	1441
Db	4776	AACGTTCGGGCAGGACTGGCAGGGAAGCCAGGCAATCTACAGATTTGTGGCACCGGGG	4835
Qy	1442	AGCGTCTCTTCGGCATGTTTGAATCTGCTGCTCTCGAGTGTGTATGACCGGGTGTG	1501
Db	4836	AGCGCCCTCCGGCATGTTTCGACTCGTCTGCTCTGTGAGTGTGTATGACCGAGGCTGTG	4895
Qy	1502	CTTGGTATGAGCTTACGCCCGCGAGACACAGTTAGGCTACGAGCATACATGAACACC	1561
Db	4896	CTTGGTATGAGCTACGCCCGCGAGACTACAGTTAGGCTACGAGCGTACATGAACACC	4955
Qy	1562	CGGGACTTCCCGTGTGCCAGACCATCTGAAATTTTGGGAGGCGTCTTTACGGGTCTCA	1621
Db	4956	CGGGGCTTCCCGTGTGCCAGACCATCTTGAAATTTTGGGAGGCGTCTTTACAGGCTCA	5015
Qy	1622	CCCATATAGACGCCCACTTCTATCCAGACAAGCAGAGTGGGGAACCTTCCCTATC	1681
Db	5016	CTCATATAGATGCCCACTTTCTATCCAGACAAGCAGAGTGGGGAACCTTCTTACC	5075
Qy	1682	TGGTAGCGTACCAAGCCACCGTGTGCGTAGAGCTCAAGCCCTCCCGCGTCCGTGGGACC	1741
Db	5076	TGGTAGGCTACCAAGCCACCGTGTGCGTAGAGGCTCAAGCCCTCCCGCATCGTGGGACC	5135
Qy	1742	AGATGTGGAAGTGTGATCCGCTCTCAAGCCCAACCTTCCATGGGCCAACAACCTCTGCTAT	1801
Db	5136	AGATGTGGAAGTGTGATTCGCTCAAGCCCAACCTTCCATGGGCCAACAACCTCTGCTAT	5195
Qy	1802	ATAGACTGGGGCTGTCCAGATGAGTCAACCTTGACGACCCAGTCAACAGTATATCA	1861
Db	5196	ACAGACTGGGGCTGTTCAGATGAAATCAACCTTGACGACCCAGTCAACCAATATATCA	5255
Qy	1862	TGACATGATGTGCGGTGACTGGAGGTGCTACAGAGTACTGGGTGCTGTTGGCGGCG	1921
Db	5256	TGACATGATGTGCGCGGACTGGAGGTGCTACAGAGCACTGGGTGCTGTTGGCGGCG	5315
Qy	1922	TTCTGGCTGTTTGGCGCGGATTTGCTATCCACAGGCTGCGGTATAGTAGGTAGGA	1981
Db	5316	TTCTGGCTGTTTGGCGCGGATTTGCTATCCACAGGCTGCGGTATAGTAGGTAGGA	5375
Qy	1982	TTCTGTTGTCGGGAAGCCGGCAATCATACCCGACAGGAGTCTCTACCGGAGTTG	2041
Db	5376	TCGTCTTGTCCGGGAAGCCGGCAATCATACCTGACAGGGAAGTCTCTACCGGAGTTG	5435

Search completed: February 27, 2005, 20:54:59  
Job time : 8691 secs

Qy 2042 ATGAAATGGAAGTGTCT 2059

Db 5436 ATGAGATGGAAGTGTCT 5453

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 27, 2005, 15:10:47 ; Search time 1053 Seconds  
(without alignments)  
11586.502 Million cell updates/sec

**Title:** US-09-930-591-1  
**Perfect score:** 2061  
**Sequence:** 1 atggcgccctatcacggccta.....atgaaatggaagagatcgtga 2061

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

. Listing first 45 summaries

Database : N\_Geneseq\_16Dec04:\*

1: geneseqn1980s:\*

2: geneseqn1990s: \*

3: geneseqn2000s:\*

4: geneseqn2001as:\*

5: geneseqn2001b9: \*  
6: geneseqn2002a8: \*

7: geneseqn2002bs:\*

8: geneseg2003as:\*

9: geneseqn2003bs:\*

10: geneseqn2003cs:

```

11: geneseqn2003ds:
12: geneseqn2004ag:

```

12: [geneseqn2004ab](#):  
13: [geneseqn2004bs](#):

.....

Pred. No. is the number of res

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query		ID	Description
		Match	Length		
1	2061	100.0	2061	6 AAD34500	Hepatitis
2	2061	100.0	2061	6 AAD31767	Hepatitis
3	2061	100.0	2061	10 RAD50868	Hepatitis
4	2061	100.0	2061	12 ADG47658	Hepatitis
5	1787	86.7	6299	4 AAF83669	HCV NS3/4
6	1786	86.7	7310	1 AAN92106	Combined
7	1786	86.7	7310	1 AAN90336	Composite
8	1786	86.7	7310	2 AAO98221	Hepatitis
9	1786	86.7	8316	3 AA75296	cDNA sequ
10	1786	86.7	9133	2 AAZ07656	Nucleotide
11	1786	86.7	9185	2 AAQ01566	Hepatitis
12	1786	86.7	9185	3 AA75297	Sense str
13	1786	86.7	9185	12 ADN35979	HCV cDNA
14	1786	86.7	9401	2 AAT13710	Hepatitis
15	1786	86.7	9401	2 AAT99981	HCV poly
16	1786	86.7	9401	2 AAV03989	HCV poly
17	1786	86.7	9401	6 AAD33043	Hepatitis
18	1786	86.7	9401	12 ADL23106	Hepatitis
19	1786	86.7	9401	13 ADR29358	Hepatitis
20	1784.4	86.6	6905	1 AAN92103	Combined

21	1784.4	86.6	8316	2	AAQ05955	AaQ05955	Hepatitis
22	1784.4	86.6	9185	2	AAQ05956	AaQ05956	Sense str
23	1784.4	86.6	9185	2	AAQ26737	AaQ26737	Nucleotid
24	1784.4	86.6	9185	2	ADF60608	Adf60608	Hepatitis
25	1784.4	86.6	9400	2	AAQ21744	AaQ21744	Compiled
26	1782.8	86.5	9185	2	AAQ00459	AaQ00459	Hepatitis
27	1782.6	86.5	5300	1	AAQ20097	AaQ20097	Hepatitis
28	1781.2	86.4	2058	6	AAQ29795	AaQ29795	HCV-1 NS3
29	1781.2	86.4	2058	12	ADCL06768	Adc06768	HCV mutan
30	1781.2	86.4	2058	12	ADL66804	Adl66804	HCV NS3/4
31	1781.2	86.4	5360	1	AAQ90327	AaQ90327	Hepatitis
32	1781.2	86.4	9618	11	AAQ933102	AaQ933102	Hepatitis
33	1779.6	86.3	2058	6	ABK15344	Abk15344	Hepatitis
34	1779.6	86.3	2058	10	ABX14410	Abx14410	DNA encod
35	1776.4	86.2	8316	12	AAQ59364	AaQ59364	HCV cdNA
36	1774.8	86.1	12980	2	AAV59364	Aav59364	Hepatitis
37	1774.8	86.1	12980	8	ACA62469	AcA62469	DNA encod
38	1774.8	86.1	16622	3	AAZ36212	Aaz36212	Nucleotid
39	1773.2	86.0	9379	2	AAQ36209	AaQ36209	Composite
40	1773.2	86.0	9646	2	AAV59361	Aav59361	Hepatitis
41	1773.2	86.0	9646	6	ABK87285	Abk87285	cDNA enco
42	1773.2	86.0	9646	8	ACA62466	AcA62466	HCV H77 c
43	1773.2	86.0	12980	6	ABK87286	Abk87286	Hepatitis
44	1765.2	85.6	9518	5	AAQ03778	AaQ03778	Hepatitis
45	1765.2	85.6	9599	2	AAQ24833	AaQ24833	Infectiou

## ALIGNMENTS

RESULT 1	
AAAD34500	
ID	AAAD34500 standard; DNA; 2061 BP.
XX	
XX	
AAAD34500;	
XX	
XX	
DT	16-JUL-2002 (first entry)
XX	
XX	Hepatitis C virus NS3/4A protein encoding DNA.
DE	
XX	Hepatitis C virus; HCV; NS3/4A protein; therapy
KW	virucide; gene; ds.
KW	
XX	
OS	Hepatitis C virus.
XX	
XX	
FH	Key
FT	CDS
FT	1. .2061
FT	/*tag= a
FT	/product= "HCV NS3/4A protein"
XX	
PN	WO200214362-A2.
XX	
XX	
PD	21-FEB-2002.
XX	
XX	
PF	15-AUG-2001; 2001WO-IB001774.
XX	
PR	17-AUG-2000; 2000US-0225767P.
PR	29-AUG-2000; 2000US-0229175P.
PR	03-NOV-2000; 2000US-00705547.
XX	
XX	
PA	(TRIP-) TRIPEP AB.
XX	
PI	Sallberg M;
XX	
XX	WPI; 2002-3399446/37.
DR	P-PSDB; AAE21837.
DR	
XX	
XX	
PT	Novel hepatitis C virus NS3/4A peptide useful
PT	absence of hepatitis C virus in a subject and
PT	for treating hepatitis C virus infection.
XX	
XX	Claim 1; Page 64-65; 90pp; English.
XX	
XX	

CC The present invention relates to novel hepatitis C virus (HCV) NS3/4A  
CC proteins and their corresponding polynucleotides. NS3/4A sequences are  
CC useful for identifying the presence or absence of HCV in a subject. They are  
CC useful for preparing a medicament used for treating or preventing HCV  
CC infection. Sequences of the invention are also used as vaccines. The  
CC present sequence is a DNA encoding HCV NS3/4A protein  
XX  
SQ Sequence 2061 BP; 427 A; 616 C; 571 G; 447 T; 0 U; 0 Other;

Query Match 100.0%; Score 2061; DB 6; Length 2061;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 2061; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGCGCTATACGGCTATGCCAGCAGACAAAGGGCGCTTTGGATGCATAATCACC 60  
DB 1 ATGGCGCTATACGGCTATGCCAGCAGACAAAGGGCGCTTTGGATGCATAATCACC 60

QY 61 AGCTTGACCGCGCGGACAAAACCAAGGTGGAGGTGAGGTTGAGATCGTGTCAACTGCT 120  
DB 61 AGCTTGACCGCGCGGACAAAACCAAGGTGGAGGTGAGGTTGAGATCGTGTCAACTGCT 120

QY 121 GCCAGACTTTCTTGGCAACTGCATTAAGCGGGTGTGTGGACTGTCTACCATGGAGCC 180  
DB 121 GCCAGACTTTCTTGGCAACTGCATTAAGCGGGTGTGTGGACTGTCTACCATGGAGCC 180

QY 181 GGAAACAGGACCATTCGTCACCTAAGGTCCTGTTATCCAGATGTACACCAATGTGGAC 240  
DB 181 GGAAACAGGACCATTCGTCACCTAAGGTCCTGTTATCCAGATGTACACCAATGTGGAC 240

QY 241 CAAGACCTCGTAGGCTGGCCGCTGCCAAGGTGCCCGCTCAATTAACACCATGCATTGC 300  
DB 241 CAAGACCTCGTAGGCTGGCCGCTGCCAAGGTGCCCGCTCAATTAACACCATGCATTGC 300

QY 301 GGCTCTCGAACCTTTACCTGTGACAGGACGCGGATGTCAATTCCTGTGCGCCGACGG 360  
DB 301 GGCTCTCGAACCTTTACCTGTGACAGGACGCGGATGTCAATTCCTGTGCGCCGACGG 360

QY 361 GGTGATGGCAGGCGAGCTGCTTTGCGCCGCGCTATCTTACTTTGAAAGCTCCTCG 420  
DB 361 GGTGATGGCAGGCGAGCTGCTTTGCGCCGCGCTATCTTACTTTGAAAGCTCCTCG 420

QY 421 GGAGGCGCTCTGCTGTGCCCGCAGGACATTCAGAGCGCGGTATGC 480  
DB 421 GGAGGCGCTCTGCTGTGCCCGCAGGACATTCAGAGCGCGGTATGC 480

QY 481 ACCGTGGAGTGGCTAAGGCGGTGGACTTCATCCCGGTAGAGCTTAGAGCAACCATG 540  
DB 481 ACCGTGGAGTGGCTAAGGCGGTGGACTTCATCCCGGTAGAGCTTAGAGCAACCATG 540

QY 541 AGGTCCCGGTGTTCTCAGACAACTCCTCCCGCAGCAGTGCCTCCAGAGCTACCAAGTG 600  
DB 541 AGGTCCCGGTGTTCTCAGACAACTCCTCCCGCAGCAGTGCCTCCAGAGCTACCAAGTG 600

QY 601 GCCCACTGCTATGCCCGCAGCGGTAAAGCAACCAAGTCCCGCGCATACGCA 660  
DB 601 GCCCACTGCTATGCCCGCAGCGGTAAAGCAACCAAGTCCCGCGCATACGCA 660

QY 661 GCTCAGGGCTACAAGTGTGTGCTCAACCCCTCGTGTGCTGCAACATGGGCTTTGGT 720  
DB 661 GCTCAGGGCTACAAGTGTGTGCTCAACCCCTCGTGTGCTGCAACATGGGCTTTGGT 720

QY 721 GCTTACATGTCAGGCGCATGGATGATGATCCTTAACATCAGGACTGGGGTGAGACAAT 780  
DB 721 GCTTACATGTCAGGCGCATGGATGATGATCCTTAACATCAGGACTGGGGTGAGACAAT 780

QY 781 ACTACTGGACCGCGATCAAGTATTCACCTAGCGCAAGTTCCTGCGCGCGGGTGT 840  
DB 781 ACTACTGGACCGCGATCAAGTATTCACCTAGCGCAAGTTCCTGCGCGCGGGTGT 840

QY 841 TCAGGGGGTGTATGACATAAATTTGTGAGAGTGCACCTCCAGGATGCAACATCC 900  
DB 841 TCAGGGGGTGTATGACATAAATTTGTGAGAGTGCACCTCCAGGATGCAACATCC 900

QY 901 ATCTTTGGGCAATTTGGCACTGTCTCTTGACCAAGCAGAGACCGCGGGCGAGACTGATGTG 960  
DB 901 ATCTTTGGGCAATTTGGCACTGTCTCTTGACCAAGCAGAGACCGCGGGCGAGACTGATGTG 960

QY 961 CTGCGCACCGCTACCGCTCCGGGCTCCGTCACCTGTGCCCCCATCTTAACATCAGGAGGTT 1020  
DB 961 CTGCGCACCGCTACCGCTCCGGGCTCCGTCACCTGTGCCCCCATCTTAACATCAGGAGGTT 1020

QY 1021 GCTCTGTCCACTACCGGAGAGATCCCTTTTATGCAAGGCTATTCCCTCTGAAGCAATT 1080  
DB 1021 GCTCTGTCCACTACCGGAGAGATCCCTTTTATGCAAGGCTATTCCCTCTGAAGCAATT 1080

QY 1081 AAGGGGGGAGACATCTCATCTTCTGCCACTCAAGAAGAGTGGACAGCTTCGCCGA 1140  
DB 1081 AAGGGGGGAGACATCTCATCTTCTGCCACTCAAGAAGAGTGGACAGCTTCGCCGA 1140

QY 1141 AAACCTGGTCGGTGGGCGTCAATGCCGTGGCTTACTACCGGGCGCTTATGTGTCCGTC 1200  
DB 1141 AAACCTGGTCGGTGGGCGTCAATGCCGTGGCTTACTACCGGGCGCTTATGTGTCCGTC 1200

QY 1201 ATCCCGACAGTGGTGAAGTGTGTCGTGGCAACTGACGCCCTCATGACCGGCTTTTACC 1260  
DB 1201 ATCCCGACAGTGGTGAAGTGTGTCGTGGCAACTGACGCCCTCATGACCGGCTTTTACC 1260

QY 1261 GCGCAGCTTCGATTTCGGTGATAGACTGCAACACGTTGTCACCAGACAGTTCGATTCAGC 1320  
DB 1261 GCGCAGCTTCGATTTCGGTGATAGACTGCAACACGTTGTCACCAGACAGTTCGATTCAGC 1320

QY 1321 CTTGACCTTACCTTACCAATTGAGACATTCAGCTTCCCGGATGCTGTCTCCCGTACT 1380  
DB 1321 CTTGACCTTACCTTACCAATTGAGACATTCAGCTTCCCGGATGCTGTCTCCCGTACT 1380

QY 1381 CAACGTCGGGGTAGGACTGGCAGAGGAGCAGGCACTACAGATTTCGTGGCACCGGG 1440  
DB 1381 CAACGTCGGGGTAGGACTGGCAGAGGAGCAGGCACTACAGATTTCGTGGCACCGGG 1440

QY 1441 GAGCGTCTTCTGGCACTGTTTGAATCTCTCTGCGAGTGTATGACGCGGGTTGT 1500  
DB 1441 GAGCGTCTTCTGGCACTGTTTGAATCTCTCTGCGAGTGTATGACGCGGGTTGT 1500

QY 1501 GCTTGATATGAGTTAGCGCCGCGAGACACAGTTAGGCTACGACATACATGACAC 1560  
DB 1501 GCTTGATATGAGTTAGCGCCGCGAGACACAGTTAGGCTACGACATACATGACAC 1560

QY 1561 CCGGACTTCCCGTGTGCCAAGACCACTTTGAATTTGGAGGGCGCTCTTTACGGGTCTC 1620  
DB 1561 CCGGACTTCCCGTGTGCCAAGACCACTTTGAATTTGGAGGGCGCTCTTTACGGGTCTC 1620

QY 1621 ACCCATAGACGCCCACTTCTCTATCCAGACAAAGCAGTGGGAAACCTTCCCTAT 1680  
DB 1621 ACCCATAGACGCCCACTTCTCTATCCAGACAAAGCAGTGGGAAACCTTCCCTAT 1680

QY 1681 CTGGTAGCGTACAGCGCAACGCTGTGCGCTAGAGCTCAAGCCCTCCCGCTCGTGGAC 1740  
DB 1681 CTGGTAGCGTACAGCGCAACGCTGTGCGCTAGAGCTCAAGCCCTCCCGCTCGTGGAC 1740

QY 1741 CAGATGTGAAGTGTGTTGATCCGCTCAAGCCCAACCTCTCAAGGGCCAAACCTCTGCTA 1800  
DB 1741 CAGATGTGAAGTGTGTTGATCCGCTCAAGCCCAACCTCTCAAGGGCCAAACCTCTGCTA 1800

QY 1801 TATAGACTGGGCGCTGCCAGAAATGAAGTCAACCTGACGCCACCCAGTACCAAGTATATC 1860  
DB 1801 TATAGACTGGGCGCTGCCAGAAATGAAGTCAACCTGACGCCACCCAGTACCAAGTATATC 1860

QY 1861 ATGACATGATGTGGCTGACCTGAGAGTGTCTCAGAGTACCTGGGTGCTCGTGGCGG 1920  
DB 1861 ATGACATGATGTGGCTGACCTGAGAGTGTCTCAGAGTACCTGGGTGCTCGTGGCGG 1920

QY 1921 GTTCTGGCTGCTTTGGCGCGTATTCGCTATCCAGAGGTGGTGTCTATAGTAGGTAGG 1980  
DB 1921 GTTCTGGCTGCTTTGGCGCGTATTCGCTATCCAGAGGTGGTGTCTATAGTAGGTAGG 1980

QY 1981 ATTGTCTTGTCCGAAAGCGGCAATCATACCGCACAGGGAAGTCTCTTACCGGGAGTTC 2040



Db	1981	ATTGTCCTTGTCCGGAAGCCGGCAATCATACCCACAGGGAAGTCCTCTTACCGGAGTTC	2040
Qy	2041	GATGAAATGGAAGAGTGCTGA	2061
Db	2041	GATGAAATGGAAGAGTGCTGA	2061
RESULT 2			
AAD31767			
ID	AAD31767	standard; DNA; 2061 BP.	
XX			
AC	AAD31767;		
XX			
DT	18-JUN-2002	(first entry)	
XX			
DE	Hepatitis C virus (HCV)	NS3/4A DNA coding region.	
XX			
KW	Hepatitis C virus; HCV infection; virucide; fungicide; antibacterial;		
KW	cytostatic; immunostimulant; vaccine; ribavirin; immune response; cancer;		
KW	ds.		
XX			
OS	Hepatitis C virus.		
XX			
EH	Key	Location/Qualifiers	
FT	CDS	1..2061	
FT		/*tag= a	
FT		/product= "HCV NS3/4A protein"	
XX			
PN	WO200213855-A2.		
XX			
PD	21-FEB-2002.		
XX			
PF	15-AUG-2001; 2001WO-IB001808.		
XX			
PR	17-AUG-2000; 2000US-0225767P.		
PR	29-AUG-2000; 2000US-0229175P.		
PR	03-NOV-2000; 2000US-00705547.		
XX			
PA	(TRIP-) TRIPEP AB.		
XX			
PI	Sallberg M, Hultgren C;		
XX			
DR	WPI; 2002-241837/29.		
DR	P-PSDB; AAE19900.		
XX			
PT	Vaccine compositions for treating and preventing disease, preferably		
PT	Hepatitis C virus infection, comprises ribavirin and antigen that has		
PT	epitope present in hepatitis C virus.		
XX			
PS	Claim 1; Page 94-95; 120pp; English.		
XX			
CC	The invention relates to a composition comprising ribavirin and an		
CC	antigen preferably non structural 3 protein (NS3)/4A fragment of		
CC	Hepatitis C virus (HCV) genome or a peptide or nucleic acid of HCV		
CC	sequence. The composition is useful for enhancing an immune response to a		
CC	Hepatitis C antigen in humans, domestic, sport or pet species and as		
CC	vaccines for treating and preventing HCV infections. The composition is		
CC	also useful for treating viral, bacterial, fungal diseases and cancer.		
CC	The present sequence is HCV NS3/4A DNA coding region		
XX			
SQ	Sequence 2061 BP; 427 A; 616 C; 571 G; 447 T; 0 U; 0 Other;		
Query Match	100.0%;	Score 2061; DB 6; Length 2061;	
Best Local Similarity	100.0%;	Pred. No. 0;	
Matches 2061; Conservative	0; Mismatches	0; Indels	0; Gaps
Qy	1	ATGCGCGCTATCAGCGCTATGCCACGACAGAGGGGCGCTTTTGGGATGCATAATCACC	60
Db	1	ATGCGCGCTATCAGCGCTATGCCACGACAGAGGGGCGCTTTTGGGATGCATAATCACC	60
Qy	61	AGCTTGTACCGCGCGGGGACAAAAACACAGGTGGAGGTTGAGTTCAGATCGTGTCAACTGCT	120

```
QY 1201 ATCCGACCACTGGTGTGCTGTGCTGCAACTGACGCCCTCATGACGGCTTTACC 1260
DB 1201 ATCCGACCACTGGTGTGCTGTGCTGCAACTGACGCCCTCATGACGGCTTTACC 1260
QY 1261 GCGACTTCGATTCGGTGATAGATGCAACACGTGTGTACCCAGACAGTCGACTTCAGC 1320
DB 1261 GCGACTTCGATTCGGTGATAGATGCAACACGTGTGTACCCAGACAGTCGACTTCAGC 1320
QY 1321 CTTGACCCCTACCTTCAACATTTGACAAATCAGCTTCCCGAGATGCTGTCTCCGTAAT 1380
DB 1321 CTTGACCCCTACCTTCAACATTTGACAAATCAGCTTCCCGAGATGCTGTCTCCGTAAT 1380
QY 1381 CAACGTCGGGTAGGACTGGCAGAGGAAGCCAGGCATCTACAGATTTGTGGACCGGG 1440
DB 1381 CAACGTCGGGTAGGACTGGCAGAGGAAGCCAGGCATCTACAGATTTGTGGACCGGG 1440
QY 1441 GAGCGTCTTCTGGCATGTTTGAATCTGCTCTGCTGCGAGTGTCTATGACGGGGTGT 1500
DB 1441 GAGCGTCTTCTGGCATGTTTGAATCTGCTCTGCTGCGAGTGTCTATGACGGGGTGT 1500
QY 1501 GCTTGGTATGAGCTTACGCCCGCGAGACACACAGTTAGGCTACGAGCATACATGACACC 1560
DB 1501 GCTTGGTATGAGCTTACGCCCGCGAGACACACAGTTAGGCTACGAGCATACATGACACC 1560
QY 1561 CCGGACTTCCGCTGTGCCAAGACCATCTTGAATTTTGGAGGGCGTCTTTACGGGTCTC 1620
DB 1561 CCGGACTTCCGCTGTGCCAAGACCATCTTGAATTTTGGAGGGCGTCTTTACGGGTCTC 1620
QY 1621 ACCACATAGACGCCCATCTTCTATCCAGACAAAGCAGAGTGGGGAAACCTTCCCTAT 1680
DB 1621 ACCACATAGACGCCCATCTTCTATCCAGACAAAGCAGAGTGGGGAAACCTTCCCTAT 1680
QY 1681 CTGTAGCTACCAAGACCGCTGTGCTAGAGCTCAAGCCCTCCCGCTCGTGGGAC 1740
DB 1681 CTGTAGCTACCAAGACCGCTGTGCTAGAGCTCAAGCCCTCCCGCTCGTGGGAC 1740
QY 1741 CAGATGTGGAAGTCTTGATCCGCTCAAGCCCAACCTCCATGGGCCAACACCTCTGCTA 1800
DB 1741 CAGATGTGGAAGTCTTGATCCGCTCAAGCCCAACCTCCATGGGCCAACACCTCTGCTA 1800
QY 1801 TATAGACTGGCGCTGTCCAGAAATGAAGTCAACCTGACGACCCAGTCACCAAGTATATC 1860
DB 1801 TATAGACTGGCGCTGTCCAGAAATGAAGTCAACCTGACGACCCAGTCACCAAGTATATC 1860
QY 1861 ATGACATGATGTCCGCTGACCTGGAGTGTGTGACAGTACCTGGGTCTCGTTGGCGGC 1920
DB 1861 ATGACATGATGTCCGCTGACCTGGAGTGTGTGACAGTACCTGGGTCTCGTTGGCGGC 1920
QY 1921 GTTCTGGCTCTTTGGCCGGTATTGCTTATCCACAGGCTGCGTGTATAGTAGGTAGG 1980
DB 1921 GTTCTGGCTCTTTGGCCGGTATTGCTTATCCACAGGCTGCGTGTATAGTAGGTAGG 1980
QY 1981 ATTGTCTTGTCCGAAAGCCGCAATATACCCGACAGGGAAGTCTCTTACCGGGATTTC 2040
DB 1981 ATTGTCTTGTCCGAAAGCCGCAATATACCCGACAGGGAAGTCTCTTACCGGGATTTC 2040
QY 2041 GATGAAATGGAAGTGTCTGA 2061
DB 2041 GATGAAATGGAAGTGTCTGA 2061
```

## RESULT 3

AAD60868  
ID AAD60868 standard; DNA; 2061 BP.

AC AAD60868;

XX 15-JAN-2004 (first entry)

DT Hepatitis C virus NS3/4A DNA.

XX Ribavirin; vaccine; immune response; infection; therapy; immunostimulant;

virucide; ds.  
Hepatitis C virus.  
Key Location/Qualifiers  
CDS 1..2061  
/tag= a  
/product= "Hepatitis C virus protein"

US2002136740-A1.

26-SEP-2002.

15-AUG-2001; 2001US-00929955.

17-AUG-2000; 2000US-0225767P.

29-AUG-2000; 2000US-0229175P.

(SALL/) SALLBERG M.

(HULT/) HULTGREN C.

Sallberg M, Hultgren C;

WPI; 2003-764978/72.

P-PSDB; ABW00351.

Vaccine compositions for treating and preventing disease, preferably hepatitis C virus infection, comprises ribavirin and antigen that has epitope present in hepatitis C virus.

Claim 1; Page 60-61; Opp; English.

The invention relates to a composition comprising ribavirin and an antigen, where the antigen is derived from a hepatitis virus. The vaccine is useful in enhancing the immune response to a hepatitis C antigen where the composition is delivered to an animal identified as requiring an enhanced immune response. The vaccine is useful in the treatment and prevention of hepatitis C infection. The present sequence is Hepatitis C virus NS3/4A DNA

Sequence 2061 BP; 427 A; 616 C; 571 G; 447 T; 0 U; 0 Other;

Query Match 100.0%; Score 2061; DB 10; Length 2061;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 2061; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGCGCCTATCAGGGCCTATGCCAGCAGACAAGGGGCTTTTGGATGCAATATCACC 60

DB 1 ATGGCGCCTATCAGGGCCTATGCCAGCAGACAAGGGGCTTTTGGATGCAATATCACC 60

QY 61 ACCTTGACCGCGCGGACAAACACAGGTGGAGGTGAGGTTTCAGATCGTGTCAACTGCT 120

DB 61 ACCTTGACCGCGCGGACAAACACAGGTGGAGGTGAGGTTTCAGATCGTGTCAACTGCT 120

QY 121 GCCCAGACTTTCTTGGCAACCTGCATTAAACGGGGTGTGTGGACTGTCTACCATGGAGCC 180

DB 121 GCCCAGACTTTCTTGGCAACCTGCATTAAACGGGGTGTGTGGACTGTCTACCATGGAGCC 180

QY 181 GGAACAAGACCATTCGCTACCTAAGGTCCTGTATCCAGATGTACACCAATGTGAC 240

DB 181 GGAACAAGACCATTCGCTACCTAAGGTCCTGTATCCAGATGTACACCAATGTGAC 240

QY 241 CAAGACCTCGTAGGCTGGCCCGCTCCCAAGGTGCGCTATTAAACACCATGCACTTGC 300

DB 241 CAAGACCTCGTAGGCTGGCCCGCTCCCAAGGTGCGCTATTAAACACCATGCACTTGC 300

QY 301 GGCTCCTCGGACCTTTTACCTGTGTACAGAGGACGCGCGATTCCTGTGTGCGCGACGG 360

DB 301 GGCTCCTCGGACCTTTTACCTGTGTACAGAGGACGCGCGATTCCTGTGTGCGCGACGG 360

QY 361 GGTGTATGCGGGGAGGCTGCTTTCGCGCGGCTATCTTACTTTAAAGGCTCTTCG 420

DB 361 GGTGTATGCGGGGAGGCTGCTTTCGCGCGGCTATCTTACTTTAAAGGCTCTTCG 420

Qy 421 GGAGCCCTCTGCTGTGCCCCCGAGGACANCCCGTAGGCATATTACAGAGCCGCGGTATGC 480  
Db 421 GGAGCCCTCTGCTGTGCCCCCGAGGACATGCCGTAGGCATATTACAGAGCCGCGGTATGC 480  
Qy 481 ACCGTGGAGTGGCTTAGGCGGTGGACTTATCCCGTAGAGCTTAGAGCAACCATG 540  
Db 481 ACCGTGGAGTGGCTTAGGCGGTGGACTTATCCCGTAGAGCTTAGAGCAACCATG 540  
Qy 541 AGGTCCCCGGTGTCTCAGACAACTCTCCACACAGAGTCCCGACAGCTACCAAGTG 600  
Db 541 AGGTCCCCGGTGTCTCAGACAACTCTCCACACAGAGTCCCGACAGCTACCAAGTG 600  
Qy 601 GCCACCTGCAATGCTCCACCGGAGGGTAAAGACACCAAGGTCCCGGCCGACATACGCA 660  
Db 601 GCCACCTGCAATGCTCCACCGGAGGGTAAAGACACCAAGGTCCCGGCCGACATACGCA 660  
Qy 661 GCTCAGGCTACAGGTGCTGCTCTCAACCCCTCCGTTGCTGCAACATGGGCTTTGGT 720  
Db 661 GCTCAGGCTACAGGTGCTGCTCTCAACCCCTCCGTTGCTGCAACATGGGCTTTGGT 720  
Qy 721 GCTTACATGTCACAGGCCCATGGATGATCTTAACATCAGACTGGGGTGAGGACAATT 780  
Db 721 GCTTACATGTCACAGGCCCATGGATGATCTTAACATCAGACTGGGGTGAGGACAATT 780  
Qy 781 ACTACTGGCAGCCGATCACTATTCACCTACCGCAAGTTCTTGGCCGAGCGGGTGT 840  
Db 781 ACTACTGGCAGCCGATCACTATTCACCTACCGCAAGTTCTTGGCCGAGCGGGTGT 840  
Qy 841 TCAGGGGTGTTATGACATAATATTTGTGACAGTGCCTCAACGATGCAACATCC 900  
Db 841 TCAGGGGTGTTATGACATAATATTTGTGACAGTGCCTCAACGATGCAACATCC 900  
Qy 901 ATCTTGGCATTGGCACTGTCTTGACCAAGCAGAGACCGGGGGGAGACTCACTGTG 960  
Db 901 ATCTTGGCATTGGCACTGTCTTGACCAAGCAGAGACCGGGGGGAGACTCACTGTG 960  
Qy 961 CTCGCCACCGTACCCCTCGGGTCCGTCATCTGTGCCCTCACTCAATCGAGAGGTT 1020  
Db 961 CTCGCCACCGTACCCCTCGGGTCCGTCATCTGTGCCCTCACTCAATCGAGAGGTT 1020  
Qy 1021 GCTCTGTCCACTACGGAGATCCCTTTTATGGCAAGGTATTCCTCTGAAGCAATT 1080  
Db 1021 GCTCTGTCCACTACGGAGATCCCTTTTATGGCAAGGTATTCCTCTGAAGCAATT 1080  
Qy 1081 AAGGGGGGAGACATCTCATCTCTGCGACCTCAAGAGAGTGGAGAGTCCGCGCA 1140  
Db 1081 AAGGGGGGAGACATCTCATCTCTGCGACCTCAAGAGAGTGGAGAGTCCGCGCA 1140  
Qy 1141 AAACGTGTGCGGTTGGGCGTCAATGCGGTGCTTACTACCGCGGCTTGATGTGTCGTC 1200  
Db 1141 AAACGTGTGCGGTTGGGCGTCAATGCGGTGCTTACTACCGCGGCTTGATGTGTCGTC 1200  
Qy 1201 ATCCCGACAGTGTGAGCTGTGCTGTGGCAATGACGCGCTCATGACCGGCTTTACC 1260  
Db 1201 ATCCCGACAGTGTGAGCTGTGCTGTGGCAATGACGCGCTCATGACCGGCTTTACC 1260  
Qy 1261 GGCGACTTCGATTCGCTGATAGCTGCAACACGCTGTGACCCAGACAGTCTCAGC 1320  
Db 1261 GGCGACTTCGATTCGCTGATAGCTGCAACACGCTGTGACCCAGACAGTCTCAGC 1320  
Qy 1321 CTTGACCTTACCTTCAACATTGAGACAATCAGCTTCCCGAGGATGCTGTCTCCCGTACT 1380  
Db 1321 CTTGACCTTACCTTCAACATTGAGACAATCAGCTTCCCGAGGATGCTGTCTCCCGTACT 1380  
Qy 1381 CAACGTGGGGTAGGACTGGCAGAGGAGACGAGGATCTACAGATTGTGGCACCGGG 1440  
Db 1381 CAACGTGGGGTAGGACTGGCAGAGGAGACGAGGATCTACAGATTGTGGCACCGGG 1440  
Qy 1441 GAGCGTCTCTTCGATGTTGACTCGTCTGCTCTGCGAGTCTATGACGCGGGTGT 1500  
Db 1441 GAGCGTCTCTTCGATGTTGACTCGTCTGCTCTGCGAGTCTATGACGCGGGTGT 1500

Qy 1501 GCTTGTATGAGCTTACGCCCGCGAGACACACAGTATTAGGCTACGAGCATATGAACACC 1560  
Db 1501 GCTTGTATGAGCTTACGCCCGCGAGACACACAGTATTAGGCTACGAGCATATGAACACC 1560  
Qy 1561 CCGGACTTCCCGTGTCCAGACCATCTTGAATTTTGGGAGGCGCTTTTACGGGTCTC 1620  
Db 1561 CCGGACTTCCCGTGTCCAGACCATCTTGAATTTTGGGAGGCGCTTTTACGGGTCTC 1620  
Qy 1621 ACCCAGATAGACGCCCACTTCTATCCAGACCAAGCAGAGTGGGGAACCTTCCCTAT 1680  
Db 1621 ACCCAGATAGACGCCCACTTCTATCCAGACCAAGCAGAGTGGGGAACCTTCCCTAT 1680  
Qy 1681 CTGCTAGCTACCAAGCACAAGTGTGCTGTAGACTCAAGCCCTCCCGCTGCTGGAC 1740  
Db 1681 CTGCTAGCTACCAAGCACAAGTGTGCTGTAGACTCAAGCCCTCCCGCTGCTGGAC 1740  
Qy 1741 CAGATGTGGAAGTGTCTTGAATTTTGGGAGGCGCTTTTACGGGTCTC 1800  
Db 1741 CAGATGTGGAAGTGTCTTGAATTTTGGGAGGCGCTTTTACGGGTCTC 1800  
Qy 1801 TATAGACTGGGCGCTGTCCAGAAATGAAGTCAACCTGACGACCCAGTCAACCAATATC 1860  
Db 1801 TATAGACTGGGCGCTGTCCAGAAATGAAGTCAACCTGACGACCCAGTCAACCAATATC 1860  
Qy 1861 ATGACATGTATGTGCGGTGACCTGGAGGTGTCAAGAGTACCTGGGTGCTGGTGGGCG 1920  
Db 1861 ATGACATGTATGTGCGGTGACCTGGAGGTGTCAAGAGTACCTGGGTGCTGGTGGGCG 1920  
Qy 1921 GTTCTGCTCTTTGGCGCGTATTTGCTATCCACAGCTGCGGTGCTATAGTAGTAGG 1980  
Db 1921 GTTCTGCTCTTTGGCGCGTATTTGCTATCCACAGCTGCGGTGCTATAGTAGTAGG 1980  
Qy 1981 ATTGTCTTTCGGAAGCCGCAATCATATCCGACAGGAAAGTCTCTACCGGAGTTC 2040  
Db 1981 ATTGTCTTTCGGAAGCCGCAATCATATCCGACAGGAAAGTCTCTACCGGAGTTC 2040  
Qy 2041 GATGAAATGGAAGTGTCTGA 2061  
Db 2041 GATGAAATGGAAGTGTCTGA 2061  
RESULT 4  
ID ADG47658 standard; DNA; 2061 BP.  
XX AC ADG47658;  
XX AC ADG47658;  
DT 11-MAR-2004 (first entry)  
XX HCV NS3/4A domain DNA.  
DE ds; gene; immunogen; hepatitis C virus; HCV infection; vaccine.  
XX Hepatitis C virus.  
XX Key Location/Qualifiers  
XX CDS 1..2061  
XX FT /\*tag= a  
XX FT /product= "NS3/4A domain"  
XX US2003206919-A1.  
XX 06-NOV-2003.  
XX 26-NOV-2002; 2002US-00307047.  
XX 17-AUG-2000; 2000US-0225767P.  
XX 29-AUG-2000; 2000US-0229175P.  
XX 15-AUG-2001; 2001US-00929955.  
XX 15-AUG-2001; 2001US-00930591.  
XX (SALL/) SALLBERG M.  
XX

PI Sallberg M;  
 XX WPI; 2004-051480/05.  
 DR P-PSDB; ADG47659.  
 XX  
 PT New purified or isolated nucleic acid useful for enhancing an immune  
 response to a hepatitis C antigen comprises specific nucleotide sequences  
 and the amino acid sequences.  
 XX  
 PS Example 1; SEQ ID NO 1; 83pp; English.  
 XX  
 CC The invention relates to a purified or isolated nucleic acid. The  
 peptides are useful as immunogens for the treatment and prevention of  
 hepatitis C virus (HCV) infection, in vaccine and immunogen compositions.  
 CC The nucleic acid and the peptide enhance an immune response to a  
 hepatitis C antigen and are potent immunogens. The present sequence is  
 CC used in the exemplification of the invention.  
 XX  
 SQ Sequence 2061 BP; 427 A; 616 C; 571 G; 447 T; 0 U; 0 Other;  
 Query Match 100.0%; Score 2061; DB 12; Length 2061;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 2061; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 0;  
 Qy 1 ATGCGGCTATCAGCGCTATGCCAGCAGACAAGGGGCCCTTTGGGATGCATAATCACC 60  
 Db 1 ATGCGGCTATCAGCGCTATGCCAGCAGACAAGGGGCCCTTTGGGATGCATAATCACC 60  
 Qy 61 AGCTTGACCGCGCGGACAAACACAGGTGAGGTGAGTTCAGATCGTGTCAACTGCT 120  
 Db 61 AGCTTGACCGCGCGGACAAACACAGGTGAGGTGAGTTCAGATCGTGTCAACTGCT 120  
 Qy 121 GCCCAGACTTCTTGGCAAACCTGCATTAAACGGGTGTTGGACTGCTACCAATGAGGC 180  
 Db 121 GCCCAGACTTCTTGGCAAACCTGCATTAAACGGGTGTTGGACTGCTACCAATGAGGC 180  
 Qy 181 GGAAACAAAGACCAATGCGTCACTAAGGTCCTGTTATCCAGATGTACCAATGTGGAC 240  
 Db 181 GGAAACAAAGACCAATGCGTCACTAAGGTCCTGTTATCCAGATGTACCAATGTGGAC 240  
 Qy 241 CAAGACCTCTAGGCTGGCGGCTCCCAAGGTGCCGCTCATTAACCAATGCACTTGC 300  
 Db 241 CAAGACCTCTAGGCTGGCGGCTCCCAAGGTGCCGCTCATTAACCAATGCACTTGC 300  
 Qy 301 GGCTCCTCGGACCTTTACCTGGTCAAGGCAACGGCATGTCACTCTGTGCGCCGACGG 360  
 Db 301 GGCTCCTCGGACCTTTACCTGGTCAAGGCAACGGCATGTCACTCTGTGCGCCGACGG 360  
 Qy 361 GGTGATGGAGGGGAGCGCTGCTTTCCGCCGCTATCTCTTACTTGAAGGCTCCTCG 420  
 Db 361 GGTGATGGAGGGGAGCGCTGCTTTCCGCCGCTATCTCTTACTTGAAGGCTCCTCG 420  
 Qy 421 GGAGGCCCTCTGTGTGCCCGCAGACATGCGTAGGCATATTCAGAGCCGGGTATGC 480  
 Db 421 GGAGGCCCTCTGTGTGCCCGCAGACATGCGTAGGCATATTCAGAGCCGGGTATGC 480  
 Qy 481 ACCCGTGGAGTGTGAAGCGGTGGACTTTCATCCCGGTAGAGCTTAGAGCAACCAATG 540  
 Db 481 ACCCGTGGAGTGTGAAGCGGTGGACTTTCATCCCGGTAGAGCTTAGAGCAACCAATG 540  
 Qy 541 AGTCCCGGTGTTCTCAGACAACTCTCCCAACGAGGTGCCCGCAGAGCTACCAAGTG 600  
 Db 541 AGTCCCGGTGTTCTCAGACAACTCTCCCAACGAGGTGCCCGCAGAGCTACCAAGTG 600  
 Qy 601 GCCCACTGCATGCTCCCAACGGCAGGTGTAGAGCAACAGTCCCGCCGATACGCA 660  
 Db 601 GCCCACTGCATGCTCCCAACGGCAGGTGTAGAGCAACAGTCCCGCCGATACGCA 660  
 Qy 661 GCTCAGGGCTACAAGTGTGTGCTCAACCCCTCCGTTGCTGCAACAAATGGCTTTGGT 720  
 Db 661 GCTCAGGGCTACAAGTGTGTGCTCAACCCCTCCGTTGCTGCAACAAATGGCTTTGGT 720  
 Qy 721 GCTTACATGTCCNAGGCCCATGGGATTGATCTTAACATCAGGACTGGGGTAGGACAATT 780

Db 721 GCTTACATGTCCNAGGCCCATGGGATTGATCTTAACATCAGGACTGGGGTAGGACAATT 780  
 Qy 781 ACTACTGGGAGCCGATCAGTATTCACCTACGGCAAGTTCTTCCGACGCGGTGT 840  
 Db 781 ACTACTGGGAGCCGATCAGTATTCACCTACGGCAAGTTCTTCCGACGCGGTGT 840  
 Qy 841 TCAGGGGTGCTTATCAGATTAATTTGTGACGAGTGCACCTCCACGATGCAATCC 900  
 Db 841 TCAGGGGTGCTTATCAGATTAATTTGTGACGAGTGCACCTCCACGATGCAATCC 900  
 Qy 901 ATCTTGGGATTTGGCACTGCTTACCAAGCAGAGACCGCGGGGCGAGACTGACTGTG 960  
 Db 901 ATCTTGGGATTTGGCACTGCTTACCAAGCAGAGACCGCGGGGCGAGACTGACTGTG 960  
 Qy 961 CTGCGCACCGGTACCCCTCGGGCTCGTCACTGTGCCCATCTTAACATCGAGGAGTT 1020  
 Db 961 CTGCGCACCGGTACCCCTCGGGCTCGTCACTGTGCCCATCTTAACATCGAGGAGTT 1020  
 Qy 1021 GCTCTGCTCACTACCGGAGAGATCCCTTTTATGGCAAGGCTATTCCTTGNAGCAATT 1080  
 Db 1021 GCTCTGCTCACTACCGGAGAGATCCCTTTTATGGCAAGGCTATTCCTTGNAGCAATT 1080  
 Qy 1081 AAGGGGGGAGAGACATCTCATCTTCTGCCACTCAAAGAAGAGTGCAGCAGCTCGCGCA 1140  
 Db 1081 AAGGGGGGAGAGACATCTCATCTTCTGCCACTCAAAGAAGAGTGCAGCAGCTCGCGCA 1140  
 Qy 1141 AAATCTGCTCGGTGGGCTCAATGCGGTGCTTACTACCGGGCTTGAATGTCCGTC 1200  
 Db 1141 AAATCTGCTCGGTGGGCTCAATGCGGTGCTTACTACCGGGCTTGAATGTCCGTC 1200  
 Qy 1201 ATCCGACCAAGTGTGACGTGTCGTCGGCAACTGACGCCCTCATGACCGCTTTACC 1260  
 Db 1201 ATCCGACCAAGTGTGACGTGTCGTCGGCAACTGACGCCCTCATGACCGCTTTACC 1260  
 Qy 1261 GCGCACTTCGATTTCGCTGATAGACTGCAACAGTGTGTCAACAGACAGTGCACCTTCAGC 1320  
 Db 1261 GCGCACTTCGATTTCGCTGATAGACTGCAACAGTGTGTCAACAGACAGTGCACCTTCAGC 1320  
 Qy 1321 CTTGACCTTACCTTACCATTTGAGACATCAAGTTCCTCCAGGATGCTGTCTCCGCTACT 1380  
 Db 1321 CTTGACCTTACCTTACCATTTGAGACATCAAGTTCCTCCAGGATGCTGTCTCCGCTACT 1380  
 Qy 1381 CAACGTTCGGGTAGAGCTGGCAGAGGGAAGCAGGACATCTACAGATTTGTGACACCGGG 1440  
 Db 1381 CAACGTTCGGGTAGAGCTGGCAGAGGGAAGCAGGACATCTACAGATTTGTGACACCGGG 1440  
 Qy 1441 GAGCGTCTTCTTGGCATGTTTGAATCGTCTGTCTCTGCGAGTGTATGACGCGGTGT 1500  
 Db 1441 GAGCGTCTTCTTGGCATGTTTGAATCGTCTGTCTCTGCGAGTGTATGACGCGGTGT 1500  
 Qy 1501 GCTTGGTATGAGCTTACGCCCGCAGACCAAGTTCAGTTCAGGATAGAGCATACATGAACACC 1560  
 Db 1501 GCTTGGTATGAGCTTACGCCCGCAGACCAAGTTCAGTTCAGGATAGAGCATACATGAACACC 1560  
 Qy 1561 CCGGACTTCCCGTGTGCAAGACCATCTTGAATTTGGAGGGCTCTTTACGGGTCTC 1620  
 Db 1561 CCGGACTTCCCGTGTGCAAGACCATCTTGAATTTGGAGGGCTCTTTACGGGTCTC 1620  
 Qy 1621 ACCCAGATAGACGCCCATCTTCTATCCAGACAAAGCAGAGTGGGGAACCTTCCCTAT 1680  
 Db 1621 ACCCAGATAGACGCCCATCTTCTATCCAGACAAAGCAGAGTGGGGAACCTTCCCTAT 1680  
 Qy 1681 CTGGTAGCGTACCAAGCCACCGTGTGCGTAGAGCTCAAGCCCTCCCGCTCGTGGAC 1740  
 Db 1681 CTGGTAGCGTACCAAGCCACCGTGTGCGTAGAGCTCAAGCCCTCCCGCTCGTGGAC 1740  
 Qy 1741 CAGATGTGGAAGTGTTCATCCGCTCAAGCCACCTCCATGGGCCAACCTCTGCTA 1800  
 Db 1741 CAGATGTGGAAGTGTTCATCCGCTCAAGCCACCTCCATGGGCCAACCTCTGCTA 1800  
 Qy 1801 TATAGACTGGGCGTGTCCAGATGAAGTCACTCCGTGACGCCACCAAGTCAACAGTATATC 1860

Db 1801 TATAGACTGGGGCTGTCCAGAAATGAAGTACACCTGACGACCCAGTCACCAAGTATATC 1860  
Qy 1861 ATGACATGATATGTCGGCTACCTGGAGTGCTCAGAGTACCTGGGTCTCGTTGGCGGC 1920  
Db 1861 ATGACATGATATGTCGGCTACCTGGAGTGCTCAGAGTACCTGGGTCTCGTTGGCGGC 1920  
Qy 1921 GTTCTGGCTGCTTTGGCGCGTATGCTATCCACAGGCTGGTGGTCATAGTAGTAGG 1980  
Db 1921 GTTCTGGCTGCTTTGGCGCGTATGCTATCCACAGGCTGGTGGTCATAGTAGTAGG 1980  
Qy 1981 ATTGCTGTGTCGGGAAGCCGGCAATCATATCCACAGGGAAGTCTCTACCGGGAGTTT 2040  
Db 1981 ATTGCTGTGTCGGGAAGCCGGCAATCATATCCACAGGGAAGTCTCTACCGGGAGTTT 2040  
Qy 2041 GATCAAAATGGAAGTGTCTGA 2061  
Db 2041 GATCAAAATGGAAGTGTCTGA 2061  
RESULT 5  
AAF83669  
ID AAF83669 standard; DNA; 6299 BP.  
XX AC AAF83669;  
XX 23-JUL-2001 (first entry)  
DE HCV NS34A ORF comprising pCMV-NS34A nucleic acid sequence.  
KW HCV; mutant; non-structural protein; NS; hepatitis C virus; mutation;  
KW catalytic domain; NS3; NS4; NS5; antiviral; vaccine; immunostimulant;  
KW immunotherapy; NS34A; ds.  
XX OS Synthetic.  
OS Hepatitis C virus.  
XX FH Key Location/Qualifiers  
FT CDS 1990..4050  
FT FT /\*tag= a  
FT FT /product= "NS34A"  
XX W0200138360-A2.  
XX 31-MAY-2001.  
XX 22-NOV-2000; 2000WO-US032326.  
XX 24-NOV-1999; 99US-0167502P.  
XX (CHIR ) CHIRON CORP.  
XX Colt D, Medina-Selby A, Selby M, Houghton M;  
XX WPI; 2001-343948/36.  
XX P-PSDB; AAB62633.  
XX Mutant non-structural (NS) Hepatitis C virus (HCV) polypeptide, useful as  
XX a vaccine against HCV, comprises a polypeptide having a mutation that  
XX functionally disrupts the catalytic domain of NS3.  
XX Disclosure; Fig 9; 340pp; English.  
XX The invention relates to an isolated mutant non-structural (NS) Hepatitis  
XX C virus (HCV) polypeptide, comprising a polypeptide having a mutation in  
XX the catalytic domain of NS3, where the mutation functionally disrupts the  
XX catalytic domain. The NS mutant polypeptides can include NS3, NS4 (NS4a  
XX and NS4b) NS5 (NS5a and NS5b) or portions thereof. The HCV polypeptide  
XX and polynucleotide (preferably DNA or a plasmid) compositions can be used  
XX in vaccines against HCV and as diagnostics. The antibodies raised against  
XX these polypeptides can also be used as diagnostics, or for passive  
XX immunotherapy. The antibodies are also useful for isolating and  
XX identifying HCV particles. The present sequence represents the nucleic  
XX acid sequence of the pCMV-NS34A comprising the ORF encoding a NS34A

CC polypeptide  
XX SQ Sequence 6299 BP; 1449 A; 1713 C; 1637 G; 1500 T; 0 U; 0 Other;  
Query Match 86.7%; Score 1787; DB 4; Length 6299;  
Best Local Similarity 91.7%; Pred. No. 0;  
Matches 1889; Conservative 0; Mismatches 170; Indels 0; Gaps 0;  
Qy 1 ATGCGCCTTATCAGCGCTATGCCAGCAGACAAAGGGGCTTTTGGGATGCATATCACC 60  
Db 1990 ATGCGCGCCATCAGCGGTACGCCACGACAGAAAGGGGCTCTCTAGGTGTCATATCACC 2049  
Qy 61 AGCTTGACCGCGCGGAGAAAAACACAGGTGAGGGTGAAGTTAGATCGTGTCACTGCT 120  
Db 2050 AGCTTAATCGCGCGGACAAAAACCAAGTGAAGGGTGAAGTTAGTGTCACTGCT 2109  
Qy 121 GCCCAGAGCTTTCTTGGCAACTGCATTAAACGGGGTGTGGAGCTGTCTACCATGGAGCC 180  
Db 2110 GCCCAAACTTCTTGGCAAGTGCATCAATGGGGTGTGTGGAGCTGTCTACCATGGGGCC 2169  
Qy 181 GGAACAAGAGACCAATTCGCTCACCTAAGGGTCTCTTTATCCAGATGTACCAATGTGGAC 240  
Db 2170 GGAACGAGGACCAATTCGCTCACCTAAGGGTCTCTTTATCCAGATGTATACCAATGTAGAC 2229  
Qy 241 CAAGACCTCTAGTGGCTGGCGCGCTCCCAAGGTGCCCGCTCATTTAAACCATGCACTTGC 300  
Db 2230 CAAGACCTTGTGGGCTGGCGCGCTTCGAAAGGTACCGGCTCATTTGACACCCCTGCACTTGC 2289  
Qy 301 GGCTCTCGACCTTTTACCTGTCAGAGGCACGCGATGTCATTCCTGTGGCGCGGAGCG 360  
Db 2290 GGCTCTCGACCTTTTACCTGTCAGAGGCACGCGATGTCATTCCTGTGGCGCGGAGCG 2349  
Qy 361 GGTGATGCGAGGGGCGAGCTGCTTTCCGCCCGGCTATCTTACTTTGAAAGGCTCTCTCG 420  
Db 2350 GGTGATGAGAGGGGCGAGCTGCTTCCGCCCGGCTATTCCTACTTTGAAAGGCTCTCTCG 2409  
Qy 421 GAGAGCCCTCTGTGTGCCCGGAGGACATGCCGTAGGCATATTTAGAGCGCGCGGTATGC 480  
Db 2410 GGGGGTCCGCTGTGTGCCCGGAGGACGCGCTGGGCTATTTAGGGCGCGCGGTGTC 2469  
Qy 481 ACCCGTGGAGTGGCTAAGGGGTGGACTTCATCCCGTAGAGAGCTTAGAGACACCATG 540  
Db 2470 ACCCGTGGAGTGGCTAAGGGGTGGACTTTATCCCTGTGGAGAACCTTAGAGACACCATG 2529  
Qy 541 AGGTCCCGGCTGTCTCAGACAACTCTCTCCCGACAGAGTCCCGGAGAGTACCAAGTG 600  
Db 2530 AGGTCCCGGCTGTCTCAGAGTAACTCTCTCCAGAGTGTGCCCGAGAGCTTCCAGTG 2589  
Qy 601 GCCACCTGCATGCTCCACCGGAGCGGTAAGAGCACCAAGGTCCCGGCGCATACGCA 660  
Db 2590 GCTCACCTCCATGCTCCACAGGCGGCAAAAGCACCAAGGTCCCGGCTGCATATGCA 2649  
Qy 661 GCTCAGGGCTACAGGTGCTGGTGTCTCAACCCCTCGTTCGTCGCAACAATGGGCTTTGGT 720  
Db 2650 GCTCAGGGCTATAAGGTGCTAGTACTCAACCCCTCTGTTGCTGCAACACTGGGCTTTGGT 2709  
Qy 721 GCTTACATGTCCAAGGCGCATGGATGATCTTAAATCAGGACTGGGGTGGAGACAATT 780  
Db 2710 GCTTACATGTCCAAGGCTCATGGATGATCTTAAATCAGGACTGGGGTGGAGACAATT 2769  
Qy 781 ACTACTGGCAGCCCGATCAGCTATTCACCTAGCGCAAGTTCCTTGGCGCGGGGTGT 840  
Db 2770 ACCACTGGCAGCCCGATCAGCTACTCCACCTAGCGCAAGTTCCTTGGCGCGGGGTGT 2829  
Qy 841 TCAGGGGCTGCTTATGACATAAATTTGTGACAGGTGCCACTCCAGGATGCAACATCC 900  
Db 2830 TCAGGGGCTGCTTATGACATAAATTTGTGACAGGTGCCACTCCAGGATGCAACATCC 2889  
Qy 901 ATCTTGGGCTTGGCAGCTGCTTTGACCAAGAGAGACCGGGGGGAGACTGACTGTG 960  
Db 2890 ATCTTGGGCTTGGCAGCTGCTTTGACCAAGAGAGACTGCGGGGGGAGACTGCTGTG 2949  
Qy 961 CTCGCCACCGCTACCCCTCCGGGCTCCGTCACCTGTGCCCCCATCTTAAATCGAGGAGTT 1020

Db 2950 CTGCCACCCGCCACCCCTCGGGCTCGCTCACTGTGCCCATCCCAACATCGAGGAGGTT 3009  
Qy 1021 GCTCTGTCCACTACCGGAGAGATCCCTTTTATGGCAAGGCTATTCCCTTTGAAGCAATT 1080  
Db 3010 GCTCTGTCCACCCGAGAGATCCCTTTTATGGCAAGGCTATTCCCTTCGAAGTAATC 3069  
Qy 1081 AAGGGGGGAGACATCTCATCTTCTGCCACTCAAAGAAGAGTCCGACGAGCTCGCCGCA 1140  
Db 3070 AAGGGGGGAGACATCTCATCTTCTGTCAATCAAAGAAGAGTCCGACGAGCTCGCCGCA 3129  
Qy 1141 AACTGTGTGGGTGGCGGTCAATAGCCGTGGCTTACTACCGCGGCTTGAATGTGCGTC 1200  
Db 3130 AAGTGTGTCAATTTGGGCATCAATAGCCGTGGCTTACTACCGCGTCTTGAACGTGCGTC 3189  
Qy 1201 ATCCCGACAGTGTGAGCTGTGCTGTGCGCAACTGACGCCCTCATGACCGGCTTTACC 1260  
Db 3190 ATCCCGACAGCGCGGATGTGCTGTGCGCAACCGATGCCCTCATGACCGGCTTACC 3249  
Qy 1261 GGGACTTCGATTCCGTGTAGACTGCAACACGCTGTGTCAACCCAGACAGTCGACTTCAGC 1320  
Db 3250 GGGACTTCGACTCGGTGTAGACTGCAATACGCTGTGTCAACCCAGACAGTCGACTTCAGC 3309  
Qy 1321 CTTGACCCCTACCTTACCAATGAGACAAATCAACGCTTCCCCAGGATGCTGTCTCCGTA 1380  
Db 3310 CTTGACCCCTACCTTACCAATGAGACAAATCAACGCTTCCCCAGGATGCTGTCTCCGTA 3369  
Qy 1381 CAACGTCGGGTAGGACTGCGAGGGAAGCCAGGCAATCAAGATTTGTGGACCGGG 1440  
Db 3370 CAACGTCGGGTAGGACTGCGAGGGAAGCCAGGCAATCAAGATTTGTGGACCGGG 3429  
Qy 1441 GAGCGTCTTCTGGCATGTTGACTGCTGTGCTCTCTGCGAGTGTATGACGGGGTGT 1500  
Db 3430 GAGCGCCCTCCGCGATGTTGACTGCTGTGCTCTCTGCGAGTGTATGACGGGGTGT 3489  
Qy 1501 GCTTGTATGAGCTTACGCCCGCGAGACCAAGTGTAGGCTACGAGATACATGAACACC 1560  
Db 3490 GCTTGTATGAGCTACGCCCGCGAGACTACAGTGTAGGCTACGAGCTACATGAACACC 3549  
Qy 1561 CCGGGACTTCCGCTGTGCCAAGACCACTTGAATTTTGGAGGGCGTCTTACGGGTCTC 1620  
Db 3550 CCGGGGCTTCCGCTGTGCCAAGACCACTTGAATTTTGGAGGGCGTCTTACGGGGCTC 3609  
Qy 1621 ACCACATAGACGCCACTTCTTATCCAGACAAAGCAGAGTGGGAAACCTTCCCTAT 1680  
Db 3610 ACTCATATAGATCCCACTTCTATCCAGACAAAGCAGAGTGGGAGAACCTTCTTAC 3669  
Qy 1681 CTGGTAGCTACCAAGCACCGTGTGGCTAGAGCTCAAGCCCTCCCGGCTGTGGGAC 1740  
Db 3670 CTGGTAGCTACCAAGCACCGTGTGGCTAGAGCTCAAGCCCTCCCGGCTGTGGGAC 3729  
Qy 1741 CAGATGTGAAGTGTGATCCGCTCTCAAGCCCAACCTCCATGGGCCACACCTCTGCTA 1800  
Db 3730 CAGATGTGAAGTGTGATTCGCCCTCAAGCCCAACCTCCATGGGCCACACCCCTGCTA 3789  
Qy 1801 TATAGACTGGGCGCTGTCCAGATGAAGTCAACCTGACGCAACCCAGTCAACAGTATATC 1860  
Db 3790 TACAGACTGGGCGCTGTTCAGATGAATCAACCTGACGCAACCCAGTCAACAAATATC 3849  
Qy 1861 ATGACATGTATGTGCGGTGACCTGAGGTGTGTCAAGAGTACCTGGGTGTGTGGTGGGC 1920  
Db 3850 ATGACATGTATGTGCGGTGACCTGAGGTGTGTCAAGAGTACCTGGGTGTGTGGTGGGC 3909  
Qy 1921 GTTCTGCTGCTTTGGCGGTATTGCTATCCACAGGTGCTGTGCTATGATAGTAGG 1980  
Db 3910 GTTCTGCTGCTTTGGCGGTATTGCTATCCACAGGTGCTGTGCTATGATAGTAGG 3969  
Qy 1981 ATTGTCTTGTCCGGAAGCGCGCAATCATACCGACAGGGAAGTCTCTACCGGAGTTC 2040  
Db 3970 GTGCTCTTGTCCGGAAGCGCGCAATCATACCTGACAGGGAAGTCTCTTACCGAGATTC 4029  
Qy 2041 GATGAATGGAAGAGTGCT 2059

Db 4030 GATGAGATGGAAGAGTGCT 4048  
RESULT 6  
AAN92106  
ID AAN92106 standard; DNA; 7310 BP.  
XX AC AAN92106;  
XX XX  
DT 09-SEP-2004 (revised)  
DT 25-MAR-2003 (revised)  
DT 02-MAR-1990 (first entry)  
XX XX  
DE Combined ORFs of the HCV cDNAs from clones K9-1 through 15e.  
XX Hepatitis C virus; HCV; non-A, non-B hepatitis; NANBH; ss.  
XX Hepatitis C virus.  
OS Unidentified.  
XX Key Location/Qualifiers  
FH CDS 3..7310  
FT /\*tag= a  
XX  
PN EP318216-A.  
XX 31-MAY-1989.  
PD  
XX 18-NOV-1988; 88EP-00310922.  
PF  
XX 18-NOV-1987; 87US-00127114.  
PR 30-DEC-1987; 87US-00139886.  
PR 26-FEB-1988; 88US-00161072.  
PR 06-MAY-1988; 88US-00191263.  
PR 26-OCT-1988; 88US-00263584.  
PR 14-NOV-1988; 88US-00271450.  
XX  
PA (CHIR ) CHIRON CORP.  
PA (CHIR ) CHIRON CORP.  
XX Houghton M, Choo QL, Kuo G;  
XX WPI; 1989-159274/22.  
DR P-PSDB; AAP92050.  
XX  
PT Purified hepatitis C virus - and associated nucleic acids and  
PT polypeptide(s).  
XX  
PS Claim 3; Fig 47-1 - 47-8; 139pp; English.  
XX  
CC It is a double-stranded nucleotide sequence of the open reading frame  
CC (ORF) (tag a) extending through clones K9-1 to 15e of hepatitis C virus  
CC (HCV) cDNA. It can be used to make oligomeric DNA hybridisation probes to  
CC detect the presence of HCV nucleic acids in samples. The polypeptide(s)  
CC it encodes could be used as immunoassay reagents and vaccines and to  
CC generate antibodies useful in diagnosis and passive immunotherapy for HCV  
CC infection/non-A, non-B hepatitis. (Updated on 25-MAR-2003 to correct PR  
CC field.) (Updated on 25-MAR-2003 to correct PI field.)  
CC  
CC Revised record issued on 09-SEP-2004 : Correction to keywords  
XX  
SQ Sequence 7310 BP; 1491 A; 2217 C; 2058 G; 1540 T; 0 U; 4 Other;  
Query Match 86.7%; Score 1786; DB 1; Length 7310;  
Best Local Similarity 91.7%; Pred. No. 0;  
Matches 1888; Conservative 0; Mismatches 170; Indels 0; Gaps 0;  
Qy 2 TGGCGCCTATACGGCTATGCCAGACGACAAAGGGGCTTTTGGGATGCATAATCACCA 61  
Db 1729 TGGCGCCCATACGGCTACGCCAGCAGACAAAGGGGCTCTCTAGGGTGCAATACCA 1788  
Qy 62 GTTTGACCGCGCGGACAAAACCGAGGTGAGGGTTCAGATCGTGTCACTGTCTG 121



1789	DB	GCCTAACTGGCCGGGACAAAAACCAAGTGAGGGGTGAGGTCAGATTGTGTCAACTGCTG	1848
122	QY	CCCAGACTTTCTTGGCAACCTGCAATTAAACGGGGTGTGTGGACTGTCTACCATGAGACCG	181
1849	DB	CCCAAACTTCTCTGCAACGTGTCATCAATGGGGTGTCTGACTGTCTACCACGGGCGG	1908
182	QY	GAACAAGGACCATTTGGCTCACCTTAAGGGTCTGTTATTCAGATGTGACACCAANTGTGGACC	241
1909	DB	GAACGAGGACCATCCGCTCAACCAAGGGTCTGTGTCATCCAGATGTATACCAATGTAGACC	1968
242	QY	AAGACCTCTGAGGTGGCCCGCTCCCAAGGTGCCGCTCATTAATAACCATGTGACATTGGC	301
1969	DB	AAGACCTTGTGGGTGGCCCGCTCCGCAAGGTAGCCGCTCATTTGACACCCCTGACATTGGC	2028
302	QY	GCTCCTCGGACCTTTTACCTGGTCAAGGACGCGCGATGTCAATTCCTGTGCGCGCAGCGG	361
2029	DB	GCTCCTCGGACCTTTTACCTGGTCAAGGACGCGCGATGTCAATTCCTGTGCGCGCAGCGG	2088
362	QY	GTGATGCGAGGGGAGCCTGCTTTGGCCCGCGGCTATCTCTTACTTTGAAGGCTCCTCGG	421
2089	DB	GTGATAGCAGGGGAGCCTGCTGTGCGCCCGCGCCATTTCCTACTTTGAAGGCTCCTCGG	2148
422	QY	GAGGCCCTCTGCTGTGCCCCGCGAGACATGCGGTAGGCATATTCAGAGCGCGCGTATGCA	481
2149	DB	GGGGTCCGCTGTTGTGCCCCGCGGGACGCGGTGGGCATATTTAGGGCCGCGGTGTGCA	2208
482	QY	CCCCTGAGGTGGCTTAAGCGCGGTGACTTTCATCCCGGTAGAGACTTAGAGACAACCATGA	541
2209	DB	CCCCTGAGGTGGCTTAAGCGCGGTGACTTTCATCCCTGTGAGAACCTTAGAGACAACCATGA	2268
542	QY	GGTCCCCGGTGTCTCTAGACAACTCTCTCCCAACAGCAGTGCCTCCAGAGCTACCAAGTGG	601
2269	DB	GGTCCCCGGTGTCTCTAGACAACTCTCTCCCAACAGTGCCTCCAGAGCTTCACAGTGG	2328
602	QY	CCACCTGTCATGCTCCACCGGAGCGGTAAAGCACCAGGTCCTCCGCGCGCATACGCAG	661
2329	DB	CTCACCTGCCATGCTCCACAGGACAGCGGCAAAAGCACCAAGGTCCTCGGCTGTCATATGCAG	2388
662	QY	CTCAGGCTCAACAGGTGCTGTGCTCAACCCCTCGGTTGCTGCAACAATGGGCTTTGGTG	721
2389	DB	CTCAGGCTATAGTGCTAGTACTCAACCCCTCTGTTGCTGCAACACTGGGCTTTGGTG	2448
722	QY	CTTACATGTCCAAGGCCCATGGGANTGCATCTTAACATCAGGACTGGGGTGAGACAATTA	781
2449	DB	CTTACATGTCCAAGGCTCATGGGATGCATCTTAACATCAGGACCGGGGTGAGAACAAATTA	2508
782	QY	CTACTGCGAGCCGATACGATATTCACCTACGGCAAGTTCCTTGCAGCGGGCGGTGTT	841
2509	DB	CCACTGGCAGCCCCCATACGCTACTCCAACCTACGGCAAGTTCCTTGGCAGCGGGGTGCT	2568
842	QY	CAGGGGTGCTTATGACATAATTAATTTGTGACAGTGCCTACTCCAGGATGCAACATCCA	901
2569	DB	CGGGGGCGCTTATGACATAATTAATTTGTGACAGTGCCTACTCCAAGATCCACATCCA	2628
902	QY	TCTTGGGCATTGGCACTGTCTTTGACCAAGCAGAGACCGCGGGGGGAGACTGACTGTGC	961
2629	DB	TCTTGGGCATCGGCACCTGTCTTTGACCAAGCAGAGACTGCGGGGGGAGACTGTTTGTGC	2688
962	QY	TCGCCACCGCTACCCCTCGGGCTCGGTCACTGTGCCCATCTCTAACATCGAGAGGGTTG	1021
2689	DB	TCGCCACCGCACCCCTCGGGGTCCGTCACTGTGCCCATCTCCCAACATCGAGAGGGTTG	2748
1022	QY	CTCTGTCCACTACCGGAGAGATCCCTTTTATGCAAGGCTATTTCCTCTTGAAGCAATTA	1081
2749	DB	CTCTGTCCACACCGGAGAGATCCCTTTTACGGCAAGGCTATCCCTCTGAAGTAAATCA	2808
1082	QY	AGGGGGGGAGACATCTCATCTCTGCGCACCTCAAGAGAGAGTGGCAGAGCTCGCGCAA	1141
2809	DB	AGGGGGGGAGACATCTCATCTTCTGTCAATCAAGAGAGAGTGGCAGAGCTCGCGCAA	2868
1142	QY	AACGTGTCGGTGGGGCGTCAATGCGGTGCTTACTACCGGGCGCTTGATGTCTCGTCA	1201
2869	DB	AGCTGTGCGATTGGGCATCAATGCGGTGCTTACTACCGGGCTTGACGCTGTCGCTCA	2928

QY	1202	TCCCGACCAAGTGTGACGTTGTGTCGTGTGCGCAATCATGACGCCCTCATGACCGGCTTTACCG	1261
DB	2929	TCCCGACCAAGCGGCGATGTTGTGTCGTGTGCGCAACCGATGCCCTCATGACCGGCTATACCG	2988
QY	1262	GCAGCTTTCGATTCGGTGATAGACTGCAACACGTGTGTACCCACAGACAGTCGACTTCAGACC	1321
DB	2989	GCAGCTTCGACTCGGTGATAGACTGCAATACGTGTGTACCCACAGACAGTCGATTCAGACC	3048
QY	1322	TTGACCCCTACCTTCAACATTTAGAGCAATACGCTTCCCAAGATGCTGTCTCCGCTACTC	1381
DB	3049	TTGACCCCTACCTTCAACATTTAGAGCAATACGCTTCCCAAGATGCTGTCTCCGCTACTC	3108
QY	1382	AACGTTCGGGTGTAGACTTGGCAGAGGAAGCCAGGCATCTACAGATTTGTGGCACCGGGGG	1441
DB	3109	AACGTTCGGGGAGGACTGGCAGGGGAGCCAGGCATCTACAGATTTGTGGCACCGGGGG	3168
QY	1442	AGCGTCTCTTGGCAGTTTGACTCGTCTGTCTCTCGAGTGTCTATGACGCGGTTGTG	1501
DB	3169	AGCGCCCTCCGGCATCTTCGACTCGTCTCGTCTCTGTGACTGTATGACGCAAGGCTGTG	3228
QY	1502	CTTGGTATGACTTAGCCCGCGAGACCAACACTCTTGAAATTTTGGAGGGCGCTCTTTACGGGTCTCA	1561
DB	3229	CTTGGTATGACTCACGCCCGCGAGACTACAGTTAGGCTACGAGCGTACATGAACACCC	3288
QY	1562	CGGACTTCCCGTGTGCAAGACCACTCTTGAAATTTTGGAGGGCGCTCTTTACGGGTCTCA	1621
DB	3289	CGGGCTTCCCGTGTGCCAGGACCATCTTGAAATTTTGGAGGGCGCTCTTTACAGGCGCTCA	3348
QY	1622	CCCACTAGAGCGGCATCTCTATCCAGACAAGCAGAGTGGGGAACACCTTCCCTATC	1681
DB	3349	CTCATATAGATGCCACTTTTCTATCCAGACAAGCAGAGTGGGGAACACCTTCCCTATCC	3408
QY	1682	TGTTAGCGTACCAAGCCACCGTGTGCGCTAGAGCTCAAGCCCTCCCGCGTGTGGGAGC	1741
DB	3409	TGTTAGCGTACCAAGCCACCGTGTGCGCTAGAGCTCAAGCCCTCCCGCATCGTGGAGCC	3468
QY	1742	AGATGTGGAAAGTGTGATCCGTCTCAAGCCCAACCTTCCATGGGCCCAACCTCTCTAT	1801
DB	3469	AGATGTGGAAAGTGTGATTCGCTCTCAAGCCCAACCTTCCATGGGCCCAACCTCTCTAT	3528
QY	1802	ATAGACTGGCGGTGTCCAGAAATGAAGTCAACCTGACGCACCCAGTCACCAAGTATATCA	1861
DB	3529	ACAGACTGGCGGTGTTCAGAAATGAATCAACCTGACGCACCCAGTCACCAAAATATCA	3588
QY	1862	TGACATGATGTTCGGCTGACCTGGAGGTGTGTACGAGTATCTGGGTGCTCGTTGGCGGGC	1921
DB	3589	TGACATGATGTTCGGCGGACCTGGAGGTGTGTACGAGTATCTGGGTGCTCGTTGGCGGGC	3648
QY	1922	TTCTGGCTGCTTTGGCGCGGTATTTGGCTATCCACAGGCTGCGTGTCTATAGTTAGGATAGGA	1981
DB	3649	TCCTGGCTGCTTTGGCGCGGTATTTGGCTGTCAACAGGCTGCGTGTCTATAGTTAGGATAGGA	3708
QY	1982	TTGTCTTGTCCGGAAGCCGGCAATCATACCCGACAGGGAAGTCTCTACCCGGGAGTTCG	2041
DB	3709	TCGTCTTGTCCGGGAGCCGGCAATCATACCTGACAGGGAAGTCTCTACCCGAGATTCG	3768
QY	2042	ATGAATGGAAGAGTGCT	2059
DB	3769	ATGAGATGGAAGAGTGCT	3786

RESULT 7	
AAAN90336	
ID	AAAN90336 standard; DNA; 7310 BP.
XX	
AC	AAAN90336;
XX	
DT	25-MAR-2003 (revised)
DT	19-JUL-2001 (revised)
DT	01-NOV-1989 (first entry)
XX	
DE	Composite hepatitis C virus (HCV) cDNA.

XX Hepatitis C virus; cDNA; clone 15e; clone k9-1; probe; vaccine; ds.  
KW Pan troglodytes.  
XX  
OS  
XX GB2212511-A.  
PN  
XX  
XX 26-JUL-1989.  
PD  
XX  
XX 18-NOV-1988; 88GB-00027024.  
XX  
XX 18-NOV-1987; 87US-00122714.  
XX  
PR 30-DEC-1987; 87US-00139886.  
PR  
PR 26-FEB-1988; 88US-00161072.  
PR  
PR 26-OCT-1988; 88US-00263584.  
XX  
XX (CHIR ) CHIRON CORP.  
PA  
XX  
XX Houghton M, Choo QL, Kuo G;  
PI  
XX  
XX WPI: 1989-215054/30.  
DR  
XX  
XX P-PSDB; AAP90288.  
DR  
XX  
XX Hepatitis C virus gene - used for prodn. of polynucleotide probes  
PT polypeptide(s) and antibodies for diagnosis, prevention and treatment of  
PT infection.  
XX  
XX  
XX Disclosure; Fig 47; 30pp; English.  
XX  
XX The sequence shows a composite hepatitis C virus (HCV) cDNA, derived by  
CC aligning clones k9-1 through 15e in 5'-3' direction. The cDNA encodes  
CC antigens which react with antibodies in patients with non-A non-B  
CC hepatitis (NANBH). The cDNA can be used to design probes, or to  
CC synthesize polypeptides, which are used to diagnose HCV-induced NANBH, to  
CC raise antibodies for immunoassay or treatment, or to produce vaccines.  
CC See also AAP90288, and AAN90303-35. (N.B. This record was resubmitted to  
CC correct errors in the sequence.) (Updated on 25-MAR-2003 to correct PR  
XX field.)  
XX  
XX Sequence 7310 BP; 1495 A; 2218 C; 2058 G; 1539 T; 0 U; 0 Other;  
SQ  
  
Query Match 86.7%; Score 1786; DB 1; Length 7310;  
Best Local Similarity 91.7%; Pred. No. 0;  
Matches 1888; Conservative 0; Mismatches 170; Indels 0; Gaps 0;  
  
QY 2 TGGCGCTATACGGCTATGCCAGCAGACAGAGGGGCTTTGGGATGCATTAATCACCA 61  
DB 1729 TGGCGCCCATACGGCGTACGCCAGCAGACAGAGGGGCTCTTAGGGTGCATTAATCACCA 1788  
  
QY 62 GCTTAGCCGCGCGGACAAACACAGGTGAGGGTGCAGATCGTGTCAACTGCTG 121  
DB 1789 GCCTAACTGCGCGGACAAACACAGGTGAGGGTGCAGATCGTGTCAACTGCTG 1848  
  
QY 122 CCAGACTTTCTTGGGCAACTGATTAACGGGGTGTGGAGTGTCTACCAATGAGAGCG 181  
DB 1849 CCCAAACCTTCTGGCAACTGATCAATGGGGTGTGCTGGAATGCTACCAAGGGGCG 1908  
  
QY 182 GAACAGGACCATTTGGTCACTTAAGGCTCTGTTATCCAGATGTACCAATGTGGAC 241  
DB 1909 GAACAGGACCATTTGGTCACTTAAGGCTCTGTTATCCAGATGTACCAATGTGGAC 1968  
  
QY 242 AAGACCTCGTAGGCTGCGCGCTCCCAAGGTGCCCGCTCATTAACACCATGCACTTGG 301  
DB 1969 AAGACCTTGTGGCTGCGCGCTCCCAAGGTAGAGCGCTCATTAACACCATGCACTTGG 2028  
  
QY 302 GCTCTCGGACCTTACCTGGTCAAGGACCGCGATGCTATCTCTGCGCGGACGGG 361  
DB 2029 GCTCTCGGACCTTACCTGGTCAAGGACCGCGATGCTATCTCTGCGCGGACGGG 2088  
  
QY 362 GTGATGCGAGGGGACGCTCTTTCGCCCGGCTATCTTACTTGAAGGCTCTCTCG 421  
DB 2089 GTGATGCGAGGGGACGCTCTTTCGCCCGGCTATCTTACTTGAAGGCTCTCTCG 2148

QY 422 GAGGCGCTCTGCTGTCGCCCGCAGGACATGCGGTAGGCATATTTAGAGCGCGGTATGCA 481  
DB 2149 GGGGTCCGCTGTTGTGCCCGCGGGGACGCGCGTGGGCATATTTAGGGCGCGGTGTGCA 2208  
  
QY 482 CCGGTGGAGTGGCTAAGGGCGGTGGACTTTCATCCCGCTAGAGAGCTTTAGAGACAACCATGA 541  
DB 2209 CCGGTGGAGTGGCTAAGGGCGGTGGACTTTCATCCCTGTGGAGAACCTTAGAGACAACCATGA 2268  
  
QY 542 GGTCCCGGTGTTCTCAGACAACTCTCCACACACAGAGTGGCCAGAGCTACCAAGTGG 601  
DB 2269 GGTCCCGGTGTTCTCAGAGTAACTCTCTCCACAGTAGTGGCCAGAGCTTCAGGTGG 2328  
  
QY 602 CCACCTGTCATGCTCCACCGCAGCGGTAAAGAGCAACAGGTCCCGCGCGCATACGAG 661  
DB 2329 CTCACTCCATGCTCCACAGGCGGCAAGACCAAGGTCCCGGTGCATATGCGAG 2388  
  
QY 662 CTCAGGGCTACAAGGTGCTGGTCTCAACCCCTCTCCGTTCGTCACAACTATGGGGTTGGTG 721  
DB 2389 CTCAGGGCTATAAGGTGCTAGTACTCAACCCCTCTGCTGCTCAACACTGGGGTTGGTG 2448  
  
QY 722 CTTACATGTCCAAGGCCCATGGGATGATCTTAACATCAGGACTGGGGTGAAGACAATTA 781  
DB 2449 CTTACATGTCCAAGGCTCATGGGATCGATCTTAACATCAGGACCGGGGTGAGAACAAATTA 2508  
  
QY 782 CTACTGCGAGCGGATACGATATTCACCTACGGCAAGTTCCTTCCGCGAGCGGGTGT 841  
DB 2509 CCACTGGCAGCCCATCACGTTACTCACCTACGGCAAGTTCCTTCCGCGAGCGGGTGT 2568  
  
QY 842 CAGGGGTGCTTATGACATAAATTTGTGACGAGTGCCACTCCACGGATGCAACATCCA 901  
DB 2569 CGGGGGCGCTTATGACATAAATTTGTGACGAGTGCCACTCCACGGATGCAACATCCA 2628  
  
QY 902 TCTTGGGCAATTTGGCACTGTCTTGACCAAGCAGAGACCGGGGGGAGACTGACTGTGC 961  
DB 2629 TCTTGGGCAATTTGGCACTGTCTTGACCAAGCAGAGACTGCGGGGGGAGACTGCTGTGC 2688  
  
QY 962 TCGCACCCCTACCCCTCGGGCTCGTCACTGTGCCCATCTTAACATCGAGGAGGTG 1021  
DB 2689 TCGCACCCCTACCCCTCGGGCTCGTCACTGTGCCCATCTTAACATCGAGGAGGTG 2748  
  
QY 1022 CTCTGTCTACTACCGGAGAGATCCCTTTTATGGCAAGCTATTCCTCTTGAAGCAATTA 1081  
DB 2749 CTCTGTCTACTACCGGAGAGATCCCTTTTATGGCAAGCTATTCCTCTTGAAGCAATTA 2808  
  
QY 1082 AGGGGGGAGACATCTCATCTTCTGCGCACTCAAGAGAGAGTGGCAGAGCTCGCGGAA 1141  
DB 2809 AGGGGGGAGACATCTCATCTTCTGCTCATTTCAAGAGAGAGTGGCAGAGCTCGCGGAA 2868  
  
QY 1142 AACTGTGTCGCTGGCGCTCAATGCGCTTACTACCGCGCTTGTGATGTGCTGCA 1201  
DB 2869 AGCTGTGCTGCTGGCGCTCAATGCGCTTACTACCGCGCTTGTGATGTGCTGCA 2928  
  
QY 1202 TCCGACCAAGTGGTGAAGTGTGCTGGCAACTGACGCGCTCATGACCGGCTTTACCG 1261  
DB 2929 TCCGACCAAGTGGTGAAGTGTGCTGGCAACTGACGCGCTCATGACCGGCTTTACCG 2988  
  
QY 1262 GGGACTTTCGATTCGGTGTAGACTGCAACAGTGTGTCAACAGAGAGTGGCAGCTTCAGCC 1321  
DB 2989 GGGACTTTCGATTCGGTGTAGACTGCAACAGTGTGTCAACAGAGAGTGGCAGCTTCAGCC 3048  
  
QY 1322 TTGACCCCTACCTTACCAATTGAGACAATCAGCTTCCCAGGATGCTGTCTCCCGTACTC 1381  
DB 3049 TTGACCCCTACCTTACCAATTGAGACAATCAGCTTCCCAGGATGCTGTCTCCCGTACTC 3108  
  
QY 1382 AACGTGGGGTAGACTGTGCGAGAGGGAAGCGAGGATCTTAAGATTTGTGGCACCGGGG 1441  
DB 3109 AACGTGGGGTAGACTGTGCGAGAGGGAAGCGAGGATCTTAAGATTTGTGGCACCGGGG 3168  
  
QY 1442 AGCGTCTTCTTGGCATGTTTGAATCTGCTGTCTGCGAGGTGCTATGACGGGGTGTG 1501  
DB 3169 AGCGTCTTCTTGGCATGTTTGAATCTGCTGTCTGCGAGGTGCTATGACGGGGTGTG 3228  
  
QY 1502 CTTGTATGAGCTTACGGCCCGCGAGACCAAGTTTGGGTAGGCTACGAGCATACATGAACCC 1561

Db 3229 CTTGTATGAGCTACGCCCGCGAGACTACAGTTAGGCTACGAGCTACATGACACCC 3288  
1562 CGGACTTCCCGTGTGCCAAGACCATCTTGAATTTGGAGGGCGCTTTTACGGGTCTCA 1621  
3289 CGGGCTTCCCGTGTGCCAAGACCATCTTGAATTTGGAGGGCGCTTTTACAGGCCTCA 3348  
1622 CCCATAGAGCGCCACCTCTTATCCAGACAAAGCAGAGTGGGAAACCTTCCCTATC 1681  
3349 CTCATATAGATGCCACCTTCTATCCAGACAAAGCAGAGTGGGAGAACCTTCTTACC 3408  
1682 TGGTAGCTTACACAGCCAGCTGTGCGCTAGAGCTCAAGCCCTCCCGGTGTTGGGACC 1741  
3409 TGGTAGCTTACACAGCCAGCTGTGCGCTAGAGCTCAAGCCCTCCCGGTGTTGGGACC 3468  
1742 AGATGTGGAAGTCTTGTATCGCTCTCAAGCCACACCTCTCATGGGCAACACCTCTGCTAT 1801  
3469 AGATGTGGAAGTCTTGTATCGCTCTCAAGCCACACCTCTCATGGGCAACACCTCTGCTAT 3528  
1802 ATAGACTGGGCGCTGTCCAGAAATGAAGTCAACCTGACCGACCCAGTCAACCAATATCA 1861  
3529 ACAGACTGGGCGCTGTTCAGAAATGAAGTCAACCTGACCGACCCAGTCAACCAATATCA 3588  
1862 TGACATGATATCGCGCTACCTGGAGTCTGTCAGAGTACCTGGGTGCTCGTTGGCGCG 1921  
3589 TGACATGATATCGCGCGCGCTGGAGTCTGTCAGAGTACCTGGGTGCTCGTTGGCGCG 3648  
1922 TTCTGGCTGCTTTGGCGCGCTATTGCCCTATCCACAGGCTGGGTGCTCATAGTAGTAGGA 1981  
3649 TCCTGGCTGCTTTGGCGCGCTATTGCCCTATCCACAGGCTGGGTGCTCATAGTAGTAGGA 3708  
1982 TTGTCTTTGTCGGAAGCGGCAATCATACCCGACAGGGAAGTCTCTACCGGAGTTTCG 2041  
3709 TCGTCTTTGTCGGAAGCGGCAATCATACCCGACAGGGAAGTCTCTACCGGAGTTTCG 3768  
2042 ATGAATGGAGAGTGCT 2059  
3769 ATGAGATGGAGAGTGCT 3786

RESULT 8  
AAQ98221  
ID AAQ98221 standard; cDNA to mRNA; 7310 BP.  
XX AC AAQ98221;  
XX AC  
DT 25-MAR-2003 (revised)  
DT 15-AUG-1996 (first entry)  
XX DE  
XX DE Hepatitis C virus clone genome.  
XX KW Hepatitis C virus; HCV; antigen; detection; diagnosis; vaccine;  
XX KW antibodies; immunoprophylaxis; sera; serum; ds.  
XX OS  
XX OS Hepatitis C virus.  
XX FN US5443965-A.  
XX PD 22-AUG-1995.  
XX PF 05-APR-1991; 91US-00681703.  
XX PR 06-APR-1990; 90US-00505611.  
XX PR 09-OCT-1990; 90US-00594854.  
XX PR  
XX (GENE-) GENELABS INC.  
XX PI Moeckli R, Reyes GR, Kim JP;  
XX PI  
XX DR WPI; 1995-302120/39.  
XX PT New nucleic acids encoding hepatitis C virus antigens - used to develop  
PT prods. for detection of HCV-infected sera and prodn. of vaccines and anti

PT -HCV antibodies.  
XX Example 4; Fig 11; 71bp; English.  
XX Hepatitis C virus (HCV) antigens can be used for detecting HCV infected  
CC sera and individuals infected with HCV. They can also be used in an anti-  
CC HCV vaccine or for the production of anti-HCV antibodies which can be  
CC used for passive immunoprophylaxis. The antigens consistently identify  
CC more HCV positive serum samples with a high degree of specificity. See  
CC AAQ98202-14 and AAQ91939-51. (Updated on 25-MAR-2003 to correct PF  
CC field.) (Updated on 25-MAR-2003 to correct PR field.)  
XX Sequence 7310 BP; 1494 A; 2217 C; 2060 G; 1539 T; 0 U; 0 Other;  
SQ  
Query Match 86.7%; Score 1786; DB 2; Length 7310;  
Best Local Similarity 91.7%; Pred. No. 0;  
Matches 1888; Conservative 0; Mismatches 170; Indels 0; Gaps 0;  
Qy 2 TGGGCGCTATCAGCGCCTATGCCAGCAGACAAAGGGCCTTTGGGATGCATATACCA 61  
Db 1729 TGGGCGCCTATCAGCGCGTACGCCAGCAGACAAAGGGCCTCTAGGGTGCATAATCA 1788  
Qy 62 GCTTACCGCGCGGACAAACCAAGTGGAGGTGAGTTCAGATCGTGTCACTGCTG 121  
Db 1789 GCCTAACTGGCGCGGACAAACCAAGTGGAGGTGAGTTCAGATTCGTCACTGCTG 1848  
Qy 122 CCCAGACTTTCTTGGCAACCTGCAATTAACGGGGTGTGTGGACTGTCTACCATGGAGCG 181  
Db 1849 CCCAAAACCTTCTGGCAACGTGATCAATGGGGTGTGTGGACTGTCTACCAAGGGCG 1908  
Qy 182 GAACAAGGACCATGTGCTACCTAAGGGTCTGTATTATCCAGATGTACACCAATGTGGAC 241  
Db 1909 GAACGAGACCATCGCGTACCCCAAGGTCTGTCTATCCAGATGTATACCAATGTAGAC 1968  
Qy 242 AAGACCTCTGAGTGGCGCGCTCCCAAGTGGCGCTCATTAACACCATGCACTTGG 301  
Db 1969 AAGACCTCTGAGTGGCGCGCTCCCAAGTGGCGCTCATTAACACCATGCACTTGG 2028  
Qy 302 GCTCTCGGACCTTTACTGTCACAGGACGCGCGATGTCTATCTGTGCGCGACGGG 361  
Db 2029 GCTCTCGGACCTTTACTGTCACAGGACGCGCGATGTCTATCTGTGCGCGCGGG 2088  
Qy 362 GTGATGGCAGGGCAGCTGCTTTGCGCCCGCGCTATCTCTTACTTGAAAGGCTCTCGG 421  
Db 2089 GTGATAGCAGGGCAGCTGCTGTGCGCGCGCGCTATCTCTTACTTGAAAGGCTCTCGG 2148  
Qy 422 GAGCCCTCTGCTGTGCGCGCGCAGGACATGCGGTAGGCATATTAGAGCCCGCGTATGCA 481  
Db 2149 GGGTCCGCTGTTGTGCGCGCGCGGACGCGGTGGGCATATTTAGGGCCGCGGTGCA 2208  
Qy 482 CCGTGGAGTGGCTAAGCGGTGACCTTCAATCCCGGTAGAGAGCTTAGAGACAAACCATGA 541  
Db 2209 CCGTGGAGTGGCTAAGCGGTGACCTTCAATCCCGGTAGAGAGCTTAGAGACAAACCATGA 2268  
Qy 542 GGTCCCGGTGTTCTCAGACAACTCTCCACACGAGTGGCGCGCGAGAGCTTACCAAGTGG 601  
Db 2269 GGTCCCGGTGTTTACCGGATAACTCTCTCTCCACAGTAGTGGCGCGCGAGAGCTTCCAGGTG 2328  
Qy 602 CCACCTGTGATGTCTCCACCGCGCAGCGGTAAAGACCAAGGTCCCGCGCGCATACGAG 661  
Db 2329 CTCACCTTCATGTCTCCACAGGACGCGGCAAAAGGTCACCAAGGTCCCGGTGCATATGCA 2388  
Qy 662 CTCAGGGCTCAAGGTGCTGTGTCTCAACCCCTCGTGTGTCGCAACAACTGGGCTTTGGTG 721  
Db 2389 CTCAGGGCTTAAGGTGCTGTGTCTCAACCCCTCTGTGTGTCGCAACAACTGGGCTTTGGTG 2448  
Qy 722 -CTTACATGTCCAAGGCCCATGGGATGATCTTAACATCAGGACTGGGGTGAAGCAATTA 781  
Db 2449 CTTACATGTCCAAGGCTCATGGGATCGATCTTAACATCAGGACCGGGGTGAGAACAAATTA 2508  
Qy 782 CTACTGGCAGCGCGATCACTGATTTCCACCTACGCGAAGTCTCTTGGCGAGCGCGGGTGT 841  
Db 2509 CCACTGGCAGCGCGATCACTGATTTCCACCTACGCGAAGTCTCTTGGCGAGCGCGGGTGT 2568



Best Local Similarity 91.7%; Pred. No. 0;			
Matches 1888; Conservative 0; Mismatches 170; Indels 0; Gaps 0;			
Qy	2	TGGGGCTTATCACGGCTTATCCAGAGAGAAAGGGGCTTTTGGGATGCAATAATCACCA	61
Db	2735	TGGCGCCATCACGGCGTAGCGCCAGCAGACAAAGGGGCTCTTAGGGTGCAATACCA	2794
Qy	62	GCTTGACCGGCGGGGAAACAAACAGGTGGAGGTGAGTTTCAGATCGTGTCAACTGCTG	121
Db	2795	GCCTAACTGGCGGGGCAAAAACCAAGTGGAGGGTGAGGTCCAGATTGTGTCAACTGCTG	2854
Qy	122	CCAGACTTCTTCCGACCTGCAATTAACGGGTGTGTGGACTGTCTTACCATGAGCGG	181
Db	2855	CCCAAACTCTTCCGCAACGTCATCAATGGGTGTGTGGACTGTCTTACCAACGGGGCGG	2914
Qy	182	GAACAAGGACCAATTTGCGTCACCTAAGGGTCCCTGTTATCCAGATGTACACCAATGTGACC	241
Db	2915	GAACGAGGACCAATCGCGTCACCAAGGGTCTGTATCCAGATGTATACCAATGTAGACC	2974
Qy	242	AAGACCTGTTAGGTGGCCCGCTCCCCAAGGTGCCGCTCAATTAACCAATGCACTTGGC	301
Db	2975	AAGACCTGTGGGTGGCCCGCTCCGCAAGGTAGCGGCTCATTTGACACCCCTGCACTTGGC	3034
Qy	302	GCTCTCGGACCTTTACTGTGTACGAGGACGCGGATGTCAATTCCTGTGGCGCGACGGG	361
Db	3035	GCTCTCGGACCTTTACTGTGTACGAGGACGCGGATGTCAATTCCTGTGGCGCGCGGG	3094
Qy	362	GTGATGGCAGGGCAGCTGCTTTCCGCCCGGCTATCTCTTACTTGAAGGGTCCCTCGG	421
Db	3095	GTGATAGCAGGGGAGCTGTGTGCCCCGGCCCATTTCTTACTTGAAGGGTCCCTCGG	3154
Qy	422	GAGCCCTCTGCTGTGCCCGCAGGACATGCCGTAGGCATATTCAGAGCCGCGGTATGCA	481
Db	3155	GGGGTCCGCTGTTGTGCCCGCGGGCAGCGGTGGGCATATTTAGGGCGCGGTGTGCA	3214
Qy	482	CCCGTGAAGTGGCTTAAGCGGTGACTTTCATCCCGGTAGAGAGCTTAGAGACACCAATGA	541
Db	3215	CCCGTGAAGTGGCTTAAGCGGTGACTTTCATCCCGGTGGAGAACCTAGAGACACCAATGA	3274
Qy	542	GGTCCCGGTGTTCTCAGACAACTCTCCCCACAGCAGATGCCCGCAGAGCTTACCAAGTGG	601
Db	3275	GGTCCCGGTGTTTACGGATTAATCTCTCTCCACAGATAGTGCCTCCAGAGCTTCCAGGTGG	3334
Qy	602	CCCACTGCAATGCTCCACCGGAGCGGTAAAGACACCAAGGTCCTCCGCGCAGATACGAG	661
Db	3335	CTCACCTCCATGCTCCACAGGAGCGGCAAAAGCACCAAGGTCCTCCGCTGCAATATGAG	3394
Qy	662	CTCAGGGCTACAAGGTGTGTGCTCAACCCCTCGTGTGCTGCAACAAATGGCTTTGGTG	721
Db	3395	CTCAGGGCTATAAGGTGTGTGCTCAACCCCTCTGTGTGCAACACATGGGCTTTGGTG	3454
Qy	722	CTTACATGTCCAAGGCCCATGGGATTTGATCTTAACATCAGGACTGGGTGAGGACAATTA	781
Db	3455	CTTACATGTCCAGGGCTCATGGGATCGATCTTAACATCAGGACCGGGTGAGACAATTA	3514
Qy	782	CTACTGGCAGCCCGATCAGTATTTCCACTACGGCAAGTTCCTTGGCGAGCGGGGTGTT	841
Db	3515	CCACTGGCAGCCCGATCAGTATTTCCACTACGGCAAGTTCCTTGGCGAGCGGGGTGCT	3574
Qy	842	CAGGGGTGCTTATGACATAATATTTGTGACGATGCGCACTCCACGGATGCAATCCA	901
Db	3575	CGGGGGCGCTTATGACATAATATTTGTGACGATGCGCACTCCACGGATGCCACATCCA	3634
Qy	902	TCTTGGGCTATGGCAGCTGCTCTTACCAAGCAGAGACCGGGGGCGGAGCTGACTGTC	961
Db	3635	TCCTTGGGCTATGGCAGCTGCTCTTACCAAGCAGAGACTGGGGGGCGGAGCTGCTGTC	3694
Qy	962	TCGCGACCGCTTACCGGGCTCCGTCAGTGTGCCCATCTTAAACATCGAGGAGTTG	1021
Db	3695	TCGCGACCGCTTACCGGGCTCCGTCAGTGTGCCCATCTTAAACATCGAGGAGTTG	3754
Qy	1022	CTCTGTCCACTACCGGAGAGATCCCTTTTATGCAAGGGTATTTCCCTTTGAACAATTA	1081

RESULT 10  
AAZ07656

Db	3755	CTCTGTCCACCAACCGGAGAGATCCCTTTTACGGCAAGGCTATCCCCCTCGAAGTATCA	3814
Qy	1082	AGGGGGGAGACATCTCATCTTCTGCCACTCAAAGAAGAAGTGCAGAGCTCCCGCGAA	1141
Db	3815	AGGGGGGAGACATCTCATCTTCTGTCTCATTCARAGAAGAAGTGCAGAGCTCCCGCGAA	3874
Qy	1142	AACGTGTCGGTTTGGGGCTCAATGCGCTGGCTTACTACCGGGGCTTGTATGTGTCGCTCA	1201
Db	3875	AGCTGTGCGATTGGGGCTCAATGCGCTGGCTTACTACCGGGGCTTGTGACGTGTCGCTCA	3934
Qy	1202	TCCGACACAGTGTGTGCTGTGCAACTGACGCGCTCATGACGCGCTTTTACCG	1261
Db	3935	TCCGACACAGGCGATGTTGTGCTGTGGCAACCGATGCCCTCATGACCGGCTATACCG	3994
Qy	1262	GCGACTTCGATTCGGTGATAGACTGCAACACGCTGTGTCACCCAGACAGTGCAGTTCAGCC	1321
Db	3995	GCGACTTCGATTCGGTGATAGACTGCAATACGTTGTGTCACCCAGACAGTGCAGTTCAGCC	4054
Qy	1322	TTGACCCCTACCTTCCACCATTTGAGACAAATCAAGCTTCCCGAGGATGCTGTCTCCGCTACTC	1381
Db	4055	TTGACCCCTACCTTCCACCATTTGAGACAAATCAAGCTTCCCGAGGATGCTGTCTCCGCTACTC	4114
Qy	1382	AACGTGCGGGTGTGAGACTGCGAGAGGGAAGCCAGGCATCTACAGATTTGTGSCACCGGGG	1441
Db	4115	AACGTGCGGGTGTGAGACTGCGAGAGGGAAGCCAGGCATCTACAGATTTGTGSCACCGGGG	4174
Qy	1442	AGCGTCTTCTGGCATGTTTGAATCTGCTGTCTGCGAGTGTATGACGCGGTTGTG	1501
Db	4175	AGCGCCCTCCGGCATGTTTGAATCTGCTGTCTGCGAGTGTATGACGCGGTTGTG	4234
Qy	1502	CTTGGTATGAGCTTACGCGCGCGAGACCAAGTTAGGTACGAGCATATACATGAACACCC	1561
Db	4235	CTTGGTATGAGCTTACGCGCGCGAGACCAAGTTAGGTACGAGCATATACATGAACACCC	4294
Qy	1562	CGGACTTCCCGGTGCGCAAGACCATCTTGAATTTTGGGAGGGGCTTTTACGGGCTCTCA	1621
Db	4295	CGGAGCTTCCCGGTGCGCAAGACCATCTTGAATTTTGGGAGGGGCTTTTACGGGCTCTCA	4354
Qy	1622	CCACATAGACGCGCCACTTCTATCCAGACAAAGCAGAGTGGGGAACCTTCCCTATC	1681
Db	4355	CTCATATAGATGCCACTTTCATCCAGACAAAGCAGAGTGGGAGAACCTTCTTACC	4414
Qy	1682	TGTTAGGCTACCAAGCCACCGTGTGCTAGAGCTCAAGCCCTTCCCGCTGCTGGGACC	1741
Db	4415	TGTTAGGCTACCAAGCCACCGTGTGCTAGGCTCAAGCCCTTCCCGCTGCTGGGACC	4474
Qy	1742	AGATGTGGAAGTGTGATCCGCTTCAAGCCACCTTCCATGGGCCAACACCTCTGCTAT	1801
Db	4475	AGATGTGGAAGTGTGATTCGCTTCAAGCCACCTTCCATGGGCCAACACCTCTGCTAT	4534
Qy	1802	ATAGACTGGGCGCTGTCCAGATGCAAGTCAACCTGACGCAACCCAGTCAACCAAGTATCA	1861
Db	4535	ACAGACTGGGCGCTGTTCAGATGCAATCAACCTGACGCAACCCAGTCAACCAATATCA	4594
Qy	1862	TGACATGTATGTCCGCTGACCTGGAGTGTGCTACAGAGTCACTGGGTGCTGTTGGCGGCG	1921
Db	4595	TGACATGTATGTCCGCGCAGCTGGAGTGTGCTACAGAGTCACTGGGTGCTGTTGGCGGCG	4654
Qy	1922	TTCTGGCTGCTTTGGCGCGCTATTCCTATCCAGAGCTGCTGGGTGCTATAGTAGTAGGA	1981
Db	4655	TCCTGGCTGCTTTGGCGCGCTATTCCTATCCAGAGCTGCTGGGTGCTATAGTAGTAGGA	4714
Qy	1982	TTGTCTTGTCCGGAAGCCGGCAATCATACCGACAGGAGTCTCTTACCGGAGTTGCG	2041
Db	4715	TCGTCTTGTCCGGAAGCCGGCAATCATACCGACAGGAGTCTCTTACCGGAGTTGCG	4774
Qy	2042	ATGAAATGGAAGTGTCT 2059	
Db	4775	ATGAGATGGAAGTGTCT 4792	





```
Db 4484 AGCTGGTCGCAATGGGCATCAATGCGGGGCTACTACCGGGCTTTGACGTGCTCCGTCA 4543
Qy 1202 TCCGACACAGTGTGCAAGTGTCTGCTGCGTGGCAACTGAGCCCTCATGACCGGCTTTACCG 1261
Db 4544 TCCGACACAGGCGGATGTTGCTGCTGCGTGGCAACCGATGCCCTCATGACCGGCTATACCG 4603
Qy 1262 GCGACTTTCGATTCGGTGATAGACTGCGCAACAGTGTGTCAACCGACAGCTGCACTTCAGCC 1321
Db 4604 GCGACTTTCGATTCGGTGATAGACTGCGCAACAGTGTGTGTCAACCGACAGCTGCACTTCAGCC 4663
Qy 1322 TTGACCTTACCTTACCATTCAGACATCAAGCTTCCCGAGGATGCTGTCTCCCGTACTC 1381
Db 4664 TTGACCTTACCTTACCATTCAGACATCAAGCTTCCCGAGGATGCTGTCTCCCGCACTC 4723
Qy 1382 AACGTCGGGGTAGGACTGGCAGAGGAGGAGCGAGCATCTACAGATTTGTGGCACCAGGGG 1441
Db 4724 AACGTCGGGGTAGGACTGGCAGAGGAGGAGCGAGCATCTACAGATTTGTGGCACCAGGGG 4783
Qy 1442 AGCGTCTTTCGGCATGTTGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1501
Db 4784 AGCGTCTTTCGGCATGTTGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 4843
Qy 1502 CTTGGTATGACTTACGCCCGCGAGACCAAGTACAGTACGCTAGGCTACGATACATGAACACC 1561
Db 4844 CTTGGTATGACTTACGCCCGCGAGACCAAGTACAGTACGCTAGGCTACGATACATGAACACC 4903
Qy 1562 CGGACTTCCCGTGTGCCAAGACCATCTTGAATTTGGAGGGGCTCTTTACGGGTCTCA 1621
Db 4904 CGGAGCTTCCCGTGTGCCAAGACCATCTTGAATTTGGAGGGGCTCTTTACAGGCTCA 4963
Qy 1622 CCCATAGACGCCACTTCTTATCCAGACAAAGCAGAGTGGGAAACCTTCCCTATC 1681
Db 4964 CTATATAGATGCCCACTTCTATCCAGACAAAGCAGAGTGGGAGAACCTTCTTACC 5023
Qy 1682 TGATAGGCTACCAAGCACCGTGTGCGTGAGCTCAAGCCCTTCCCGCTGCTGGGACC 1741
Db 5024 TGATAGGCTACCAAGCACCGTGTGCGTGAGCTCAAGCCCTTCCCGCTGCTGGGACC 5083
Qy 1742 AGATGTGGAAGTGTGATCGCTCTCAAGCCACCTTCCATGGGCCCAACACCTCTGCTAT 1801
Db 5084 AGATGTGGAAGTGTGATCGCTCTCAAGCCACCTTCCATGGGCCCAACACCTCTGCTAT 5143
Qy 1802 ATAGCTGGGCGCTGCCAGATGAAGTCAACCTTGACGCCACCCAGTCAACCAAGTATATCA 1861
Db 5144 ACAGCTGGGCGCTGTTTCAAGATCAAAATCAACCTTGACGCCACCCAGTCAACCAATATCA 5203
Qy 1862 TGACATGATGTGCGCTGACCTGAGGTGCTCAGAGTACCTGGGTGCTGTTGGCGGCG 1921
Db 5204 TGACATGATGTGCGCGGACCTGAGGTGCTGCTCAGAGTACCTGGGTGCTGTTGGCGGCG 5263
Qy 1922 TTCTGGCTGCTTTGGCGCGGATTTGCTTATCCAGAGCTGCTGCTCATAGTAGGTAGGA 1981
Db 5264 TCCTGGCTGCTTTGGCGCGGATTTGCTTATCCAGAGCTGCTGCTCATAGTAGGTAGGA 5323
Qy 1982 TTGCTTTGTCCGAAAGCCGCAATCATACCCGACAGGAGTCTCTACCGGAGTTGCG 2041
Db 5324 TCGCTTTGTCCGAAAGCCGCAATCATACCTGACAGGAGTCTCTACCGAGTTGCG 5383
Qy 2042 ATGAATGGAAGTGTCT 2059
Db 5384 ATGAGATGGAAGTGTCT 5401
```

## RESULT 11

AAQ10566 standard; DNA; 9185 BP.

XX ID

XX AC

XX AC

XX AC

DT 25-MAR-2003

XX (revised)

DT 29-APR-1991

XX (first entry)

XX XX

DE Hepatitis C virus strain 1 DNA.

XX Hepatitis C virus; HCV-1; non-A, non-B hepatitis; HCV antigen;

KW viral infections; ss.

XX Hepatitis C virus.

OS Hepatitis C virus.

XX EP414475-A.

XX PN 27-FEB-1991.

XX PD 21-AUG-1990; 90EP-00309120.

XX PF 25-AUG-1989; 89US-00398667.

XX PR (CHIR ) CHIRON CORP.

XX PA Weiner AJ, Steimer KS;

XX PI WPI; 1991-059670/09.

XX DR Cell lines infected with hepatitis C virus - are used as source of

XX PT antigens for detection of HCV antibodies, for vaccines, and for screening

XX PT anti-viral agents.

XX PS Disclosure; Fig 1; 24pp; English.

XX CC This is a hepatitis C virus (HCV) composite cDNA sequence, deduced using

XX CC overlapping clones. A compsn. contg. the antigenic protein encoded by

XX CC this sequence is useful for detecting anti-HCV anti- bodies (Abs) and for

XX CC screening an agent which inhibits HCV replic- ation. A cell line infected

XX CC with this virus can be used as a source of antigens. The antigen is

XX CC useful for preparing vaccines for treating viral infections. See also

XX CC AAQ10567. (Updated on 25-MAR-2003 to correct PA field.)

XX SQ Sequence 9185 BP; 1849 A; 2790 C; 2608 G; 1938 T; 0 U; 0 Other;

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

```
QY 482 CCCGTGAGTGGCTAAGCGCGTGAATTTCATCCCGCTAGAGAGCTTAGAGCAACCATGA 541
DB 3876 CCCGTGAGTGGCTAAGCGCGTGGACTTTATCCCTGTGGAGAACTTAGAGCAACCATGA 3935
QY 542 GGTCCCCGGTGTCTCAGACAACTCTCCCAACAGCAGTGCCTCCAGAGCTACCAAGTGG 601
DB 3936 GGTCCCCGGTGTCTCAGAGTAACCTCTCTCCACAGTAGTGCCTCCAGAGCTTCAGGTGG 3995
QY 602 CCACCTCGCATGCTCCCAACGGCAGCGGTAAAGAGCAACAAAGTCCCGGCGCATACGCAG 661
DB 3996 CTCACTCCATGCTCCCAACAGGAGCGGCAAGAGCACAAAGTCCCGGCTGCATATGCAG 4055
QY 662 CTAGGGCTCAAGGTGCTGTGCTCAACCCCTCCGTTGCTGCAACAATGGGCTTTGGTG 721
DB 4056 CTAGGGCTCAAGGTGCTGTGCTCAACCCCTCTGTTGCTGCAACAACCTGGGCTTTGGTG 4115
QY 722 CTTCATGTCGAAGGCCATGGGATGATCTTAACATCAGGACTGGGGTGAGGACAATTA 781
DB 4116 CTTCATGTCGAAGGCTCATGGGATCGATCTTAACATCAGGACCGGGGTGAGAACAAATTA 4175
QY 782 CTACTGCGAGCCCGATCACGTTATTCACCTACGCAAGTTCCCTTGGCGAGCGGGGTGTT 841
DB 4176 CCACTGCGAGCCCGATCACGTTATTCACCTACGCAAGTTCCCTTGGCGAGCGGGGTGCT 4235
QY 842 CAGGGGTGCTTATGACATAAATATTTGTGACGAGTGCCCACTCCACGATGCAACATCCA 901
DB 4236 CGGGGGCGCTTATGACATAAATATTTGTGACGAGTGCCCACTCCACGATGCCACATCCA 4295
QY 902 TCTTGGGCATGCGACTGCTCTTGACCAAGCAGAGACCGGGGGCGAGACTGACTGTGC 961
DB 4296 TCTTGGGCATGCGACTGCTCTTGACCAAGCAGAGACTGCGGGGGCGAGACTGCTGTGTC 4355
QY 962 TCGCCACCGCTACCCCTCCGGGCTCCGTCACTGTGCCCCCATCTTAACATCGAGAGGTTG 1021
DB 4356 TCGCCACCGCCACCCCTCCGGGCTCCGTCACTGTGCCCCCATCTTAACATCGAGAGGTTG 4415
QY 1022 CTCTGTCCACTACCGGAGAGATCCCTTTTATGGCAAGGCTATTCCTCCCTTTGAAGCAATTA 1081
DB 4416 CTCTGTCCACTACCGGAGAGATCCCTTTTATGGCAAGGCTATTCCTCCCTTTGAAGTAATCA 4475
QY 1082 AGGGGGGAGACATCTCATCTTCTGCACTCAAGAGAGTGGAGAGTGGAGCTGCGGCAA 1141
DB 4476 AGGGGGGAGACATCTCATCTTCTGCACTCAAGAGAGTGGAGAGTGGAGTGGAGTGGAG 4535
QY 1142 AACTGTGCGGTGGGGCGTCAATGCCGTGCTTACTACCGCGCTTGTGATGTGCGTCA 1201
DB 4536 AGCTGTGCGATTGGGCATCAATGCCGTGCTTACTACCGCGCTTGTGATGTGCGTCA 4595
QY 1202 TCCGCAACAGTGTGAGTGTGCTGCTGGGCAACTGACGCGCTCATGACCGGCTTTACCG 1261
DB 4596 TCCGCAACAGCGCGATGTTGTGCTGCTGGCAACCGATGCCCTCATGACCGGCTATACCG 4655
QY 1262 GCGACTTCGATTCGCTGATAGACTGCAACAGTGTGCTACCCAGAGTGGAGTGGAGTGGAG 1321
DB 4656 GCGACTTCGATTCGCTGATAGACTGCAACAGTGTGCTACCCAGAGTGGAGTGGAGTGGAG 4715
QY 1322 TTGACCCCTACTCTTCAACATTTAGACCAATCACTGCTTCCCAAGGATGCTGCTCCCGTACTC 1381
DB 4716 TTGACCCCTACTCTTCAACATTTAGACCAATCACTGCTTCCCAAGGATGCTGCTCCCGTACTC 4775
QY 1382 AACGTGGGGTAGGACTGGCAGAGGGAAGCCAGGCACTCTACAGATTGTGGCACCGGGGG 1441
DB 4776 AACGTGGGGTAGGACTGGCAGAGGGAAGCCAGGCACTCTACAGATTGTGGCACCGGGGG 4835
QY 1442 AGCGTCTTCTGCACTGTTGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1501
DB 4836 AGCGTCTTCTGCACTGTTGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 4895
QY 1502 CTTTGTATGAGCTTACGCCCGCGAGACCACTAGGCTTACGAGCATATGAAACACCC 1561
DB 4896 CTTTGTATGAGCTTACGCCCGCGAGACCACTAGGCTTACGAGCATATGAAACACCC 4955
QY 1562 CGGGACTTCCCGTGTGCCAAGACCATCTTGAAATTTTGGGAGGCGCTCTTTACGGGCTCA 1621
```

```
DB 4956 CGGGGCTTCCCGTGTGCCAGGACCATCTTGAAATTTTGGAGGCGCTTTTACAGGCTCA 5015
QY 1622 CCACATAGACGCCCACTTCTTATCCAGACAAAGCAGAGTGGGAAACCTTCCCTATC 1681
DB 5016 CTATATAGATGCCCACTTCTTATCCAGACAAAGCAGAGTGGGAGAACCTTCTTACC 5075
QY 1682 TGGTAGCTACCAAGCACCGTGTGCTAGAGCTCAAGCCCCCTCCCGTCTGTTGGGACC 1741
DB 5076 TGGTAGCTACCAAGCACCGTGTGCTAGAGCTCAAGCCCCCTCCCGTCTGTTGGGACC 5135
QY 1742 AGATGTGGAAGTCTTGTATCCGCTCTCAAGCCCCCACTTCCATGGGCAACACTTCTCTAT 1801
DB 5136 AGATGTGGAAGTCTTGTATCCGCTCTCAAGCCCCCACTTCCATGGGCAACACTTCTCTAT 5195
QY 1802 ATAGACTGGGCGCTGTCCAGATGAAGTCACTGAGCAGCACCAGTCAACCAAGTATATCA 1861
DB 5196 ACAGACTGGGCGCTGTTCAGAAATGAATCACTGACGACCCAGTCAACCAATATATCA 5255
QY 1862 TGACATGTATGTGCGGTGACCTGGAGGTGCTCACAGTACCTGGGTGCTGTTGGCGGCG 1921
DB 5256 TGACATGTATGTGCGGTGACCTGGAGGTGCTCACAGTACCTGGGTGCTGTTGGCGGCG 5315
QY 1922 TTCTGCTGCTTGTGGCGGCTATTTGCTATPCACAGGCTGCGTGTGCTATAGTAGTAGGA 1981
DB 5316 TCTGCTGCTTGTGGCGGCTATTTGCTGCTCAACAGGCTGCGTGTGCTATAGTAGTAGGA 5375
QY 1982 TTGCTCTTGTCCGGAAGCGGCAATCATACCCGACAGGGAAGTCTCTACCGGAGTTCG 2041
DB 5376 TCGTCTTGTCCGGAAGCGGCAATCATACCTGACAGGGAAGTCTCTACCGGAGTTCG 5435
QY 2042 ATGAAATGGAAGTGTCT 2059
DB 5436 ATGAGATGGAAGTGTCT 5453

RESULT 12
AAA75297
ID AAA75297 standard; cDNA; 9185 BP.
XX
AC AAA75297;
XX
DT 15-JAN-2001 (first entry)
XX
Sense strand of HCV encoding a polyprotein.
XX
KW Hepatitis C virus; HCV; antisense polynucleotide; polyprotein;
XX viral infectivity; viral replication; ds.
XX
OS Hepatitis C virus.
XX
FH Key
CDS 320..9184
FT /*tag= a
FT /note= "partial sequence; no termination codon given"
XX
EP1034785-A2.
PN
PD 13-SEP-2000.
XX
PF 16-MAR-1990; 2000EP-00109602.
XX
PR 17-MAR-1989; 89US-00325338.
PR 20-APR-1989; 89US-00341334.
PR 18-MAY-1989; 89US-00355002.
PR 16-MAR-1990; 90EP-00302866.
XX
PA (CHIR ) CHIRON CORP.
XX
PI Houghton M, Choo Q, Kuo G;
XX
DR WPI; 2000-566891/53.
P-PSDB; AAB18541.
```

XX Novel composition comprising a hepatitis C virus antisense polynucleotide  
PT which is complementary to or corresponds to a sense strand of the virus  
PT genome, and selectively hybridizes to it.

XX Example; Fig 17; 75pp; English.

CC The specification describes a pharmaceutical composition which comprises  
CC a hepatitis C virus (HCV) antisense polynucleotide. The HCV is  
CC characterized by a positive stranded RNA genome which has 40% homology at  
CC the polypeptide level to a HCV polyprotein. The antisense polynucleotide  
CC binds to cellular polynucleotides which enhance and/or are required for  
CC viral infectivity, replicative ability or chronicity. The antisense  
CC polynucleotides may also be designed to bind with high specificity, to be  
CC of increased stability, to be stable and to have low toxicity. The  
CC composition also comprises an agent which causes viral RNA to be  
CC inactive. The composition is used for preventing HCV replication in a  
CC system. The present sequence represents a novel HCV cDNA sequence, which  
CC is used in the course of the invention

XX Sequence 9185 BP; 1849 A; 2790 C; 2608 G; 1938 T; 0 U; 0 Other;

Query Match	86.7%	Score 1786;	DB 3;	Length 9185;
Best Local Similarity	91.7%;	Pred. No. 0;		
Matches 1888;	Conservative 0;	Mismatches 170;	Indels 0;	Gaps 0;
Qy 2	TGGCGCCTATCACGGCCTATGCCAGCAGACAAGGGGCTTTTGGGATGCAATATCACCA	61		
Db 3396	TGGCGCCTATCACGGCCTATGCCAGCAGACAAGGGGCTTTTGGGATGCAATATCACCA	3455		
Qy 62	GCTTGACCGGCGGGGCAAAAACACAGATGAGAGGTTCAGATCGTGTCAATCGTGTG	121		
Db 3456	GCCTAACTGGCGGGGCAAAAACACAGATGAGAGGTTCAGATCGTGTCAATCGTGTG	3515		
Qy 122	CCAGAGCTTCTTGGCACTGATTAACGGGCTGTGTGACACTGTCTACCATGGAGCG	181		
Db 3516	CCCAACCTTCTTGGCACTGATTAACGGGCTGTGTGACACTGTCTACCATGGAGCG	3575		
Qy 182	GAAACAGGACCATTCGCTCACCTAAGGCTCTGTATTCAGATGTACACCAATGTGAC	241		
Db 3576	GAAACAGGACCATTCGCTCACCTAAGGCTCTGTATTCAGATGTACACCAATGTGAC	3635		
Qy 242	AAGACCTGATAGGTGGCCGCTCCCAAGGTGCCGCTCATTAACCAATGCACTTGG	301		
Db 3636	AAGACCTGATAGGTGGCCGCTCCCAAGGTGGCCGCTCATTAACCAATGCACTTGG	3695		
Qy 302	GCTCCTCGGACCTTACCTGTGTACAGGACAGCGGATGTCTTCTGTGGCGGACGG	361		
Db 3696	GCTCCTCGGACCTTACCTGTGTGTACAGGACAGCGGATGTCTTCTGTGGCGGACGG	3755		
Qy 362	GTGATGGCAGGGCAGCTGCTTTTCGCCCGGCTTCTTCTTCTTCTTCTTCTTCTTCT	421		
Db 3756	GTGATGAGGGGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	3815		
Qy 422	GAGGCCCTCTGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	481		
Db 3816	GAGGCCCTCTGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	3875		
Qy 482	CCCGTGAGTGGCTTAAGGCGGTGACCTTATCCCGGTAGAGAGCTTATAGACAAATGA	541		
Db 3876	CCCGTGAGTGGCTTAAGGCGGTGACCTTATCCCGGTAGAGAGCTTATAGACAAATGA	3935		
Qy 542	GGTCCCGGCTGCTCTACAGCACTTCTCCCGGCTGCTGCTGCTGCTGCTGCTGCTGCT	601		
Db 3936	GGTCCCGGCTGCTCTACAGCACTTCTCCCGGCTGCTGCTGCTGCTGCTGCTGCTGCT	3995		
Qy 602	CCCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	661		
Db 3996	CTCACTCCATGCTCCCAAGGACAGCGGCAAAAGCACTGCTGCTGCTGCTGCTGCTG	4055		
Qy 662	CTCAGGCTCAAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	721		
Db 4056	CTCAGGCTCAAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	4115		

Qy 722	CTTACATGTCCAAAGCCCATGGGATTCATCTTAACATCAGGACTGGGTGAGGCAATTA	781
Db 4116	CTTACATGTCCAAAGCCCATGGGATTCATCTTAACATCAGGACTGGGTGAGGCAATTA	4175
Qy 782	CTACTGGCAGCCCATCAGTATTCACCTACGCAAGTTCCTTCCGACCGGGGTGT	841
Db 4176	CCAATGGCAGCCCATCAGTATTCACCTACGCAAGTTCCTTCCGACCGGGGTGT	4235
Qy 842	CAGGGGTGCTTATGACATATAATTTGTACGAGTGCCTCCACGAGTCAACATCCA	901
Db 4236	CGGGGGGCTTATGACATATAATTTGTACGAGTGCCTCCACGAGTCAACATCCA	4295
Qy 902	TCTTGGCATTGGCACTGTCTTCAACAGCAGAGACCGGGGGCGGAGACTGCTGTC	961
Db 4296	TCTTGGCATTGGCACTGTCTTCAACAGCAGAGACTGCGGGGGCGGAGACTGCTGTC	4355
Qy 962	TCGCCACCGTACCCCTCCGGGCTCGTCACTGTGCGGCATCTTAACATCAGAGAGTTG	1021
Db 4356	TCGCCACCGTACCCCTCCGGGCTCGTCACTGTGCGGCATCTTAACATCAGAGAGTTG	4415
Qy 1022	CTCTGTCCACTACCGGAGAGATCCCTTTTATGCAAGGCTATTCCTTCCGCAATTA	1081
Db 4416	CTCTGTCCACTACCGGAGAGATCCCTTTTATGCAAGGCTATTCCTTCCGCAATTA	4475
Qy 1082	AGGGGGGAGACATCTCTCTTGTCCACTCAAAAGAAAGTGCAGAGCTCGCCGCA	1141
Db 4476	AGGGGGGAGACATCTCTCTTGTCTTCAAAAGAAAGTGCAGAGCTCGCCGCA	4535
Qy 1142	AATGTGTGGGTGGGCGTCAATGCGGTCTTATACCGGGCTTGTGATGTGCGTCA	1201
Db 4536	AGTGTGTGGGTGGGCGTCAATGCGGTCTTATACCGGGCTTGTGATGTGCGTCA	4595
Qy 1202	TCCGACACGAGTGTGAGTGTGCTGCTGCACTGACGCGCTCATGACCGGCTTACCG	1261
Db 4596	TCCGACACGAGTGTGAGTGTGCTGCTGCACTGACGCGCTCATGACCGGCTTACCG	4655
Qy 1262	GCGACTTCGATTCGCTGATAGACTGCAACACGCTGTGTCAACAGACAGTGCATTCAG	1321
Db 4656	GCGACTTCGATTCGCTGATAGACTGCAACACGCTGTGTCAACAGACAGTGCATTCAG	4715
Qy 1322	TTGACCTTACCTTACCACTGAGACATCACTGCTTCCCGAGATGCTGTCTCCGTA	1381
Db 4716	TTGACCTTACCTTACCACTGAGACATCACTGCTTCCCGAGATGCTGTCTCCCGTA	4775
Qy 1382	AAGTCTGGGTAGGACTGGCAGAGGAGAGCCAGGATCTACAGATTTGTGACACCGGG	1441
Db 4776	AAGTCTGGGTAGGACTGGCAGAGGAGAGCCAGGATCTACAGATTTGTGACACCGGG	4835
Qy 1442	AGCGTCTTCTGCGCATGTTTGAATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1501
Db 4836	AGCGTCTTCTGCGCATGTTTGAATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	4895
Qy 1502	CTTTGATAGACTTACCGCCCGGAGACCAAGTTCAGTTCAGGATCAATGACACCC	1561
Db 4896	CTTTGATAGACTTACCGCCCGGAGACTACAGTTAGGCTACGAGCGTACATGACACCC	4955
Qy 1562	CGGACTTCTCCGCTGTCAGACCAATCTTGAATTTTGGGAGGCGCTTTTACGGGTCTCA	1621
Db 4956	CGGAGCTTCTCCGCTGTCAGACCAATCTTGAATTTTGGGAGGCGCTTTTACGGGTCTCA	5015
Qy 1622	CCACATAGACGCGCTTCTCTATCCAGCAAGCAGAGTGGGGAACCTTCCCTATC	1681
Db 5016	CTCATATAGATGCCCTTCTTATCCAGCAAGCAGAGTGGGGAACCTTCCCTATC	5075
Qy 1682	TGCTAGGCTACCAAGCAGCGCTGTGCGCTAGAGCTCAAGCCCTTCCCGCTCGTGGAC	1741
Db 5076	TGCTAGGCTACCAAGCAGCGCTGTGCGCTAGAGCTCAAGCCCTTCCCGCTCGTGGAC	5135
Qy 1742	AGATGGAAGTGTGATCCGCTCAAGCCCACTTCCATGGGCCCAACCTCTGCTAT	1801
Db 5136	AGATGGAAGTGTGATCCGCTCAAGCCCACTTCCATGGGCCCAACCTCTGCTAT	5195

QY 1802 ATAGACTGGCGGTGTCAGAAATGAAGTCAACCTGACGACACCCAGTCAACCAAGTATATCA 1861  
DB |||||  
5196 ACAGACTGGCGGTGTTCAAGATGAATCACTGACGACCCAGTCAACCAATACATCA 5255  
QY 1862 TGACATGATATGTCGGCTGACCTGGAGTGTCTACAGATACCTGGGTGCTGTTGGCGGG 1921  
DB |||||  
5256 TGACATGATATGTCGGCGGACCTGGAGTGTCTACAGACACCTGGGTGCTGTTGGCGGG 5315  
QY 1922 TTCTGGCTGCTTTGGCGGCTATTGCTATCCACAGCTGCGTGGTCAATAGTAGGTAGGA 1981  
DB |||||  
5316 TCTGGCTGCTTTGGCGGCTATTGCTGTCAACAGCTGCGTGGTCAATAGTAGGTAGGA 5375  
QY 1982 TTGCTTTGTCGGGAACCGCGCAATCATACCGACAGGGAAGTCTCTACCGGAGTTG 2041  
DB |||||  
5376 TCGTCTTTGTCGGGAACCGCGCAATCATACCTGACAGGGAAGTCTCTACCGAGAGTTG 5435  
QY 2042 ATGAAATGGAAGTGTCT 2059  
DB |||||  
5436 ATGAGATGGAAGTGTCT 5453

## RESULT 13

ADN35979

ID ADN35979 standard; cDNA; 9185 BP.

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

```
Db 4476 AGGGGGGAGACATCTCATCTTCTGTCAATCAAGAAGAGTGCAGAACTCCCGCAA 4535
Qy 1142 AACTGGTCCGCTTGGGCGTCAATGCGGCTTACTACCGGCGCTTGATGTGCCGCA 1201
Db 4536 AGCTGGTCCGATTTGGGCATCAATGCGGCTTACTACCGGCGCTTGATGTGCCGCA 4595
Qy 1202 TCCGCAACAGTGTGTGAGTGTGCTGCTGCGGCAACTGAGCCCTCATGACCGGCTTTACCG 1261
Db 4596 TCCGCAACAGCGGCGATTTGCTGCTGCGGCAACCGATGCCCTCATGACCGGCTATACCG 4655
Qy 1262 GCGACTTCGATTCGGTGTAGACTGCAACAGTGTGTGTCACCGACAGTCACTTCAGCC 1321
Db 4656 GCGACTTCGATTCGGTGTAGACTGCAATAGTGTGTGTCACCGACAGTCACTTCAGCC 4715
Qy 1322 TTGACCCCTTACCTTCAACATTGAGACAATCAGCGCTTCCCGAGATGCTGTCTCCCGTACTC 1381
Db 4716 TTGACCCCTTACCTTCAACATTGAGACAATCAGCGCTTCCCGAGATGCTGTCTCCCGACTC 4775
Qy 1382 AACGTCGGGGTAGGACTGGCAGAGAGGAAAGCCAGGCACTACAGATTTGTGGCACCGGGGG 1441
Db 4776 AACGTCGGGGCAGACTGGCAGAGGGGAAAGCCAGGCACTACAGATTTGTGGCACCGGGGG 4835
Qy 1442 AGCGTCTTCTGGCATGTTGACTCGTCTGCTCTGCGAGTGTATGACCGGGTTGTG 1501
Db 4836 AGCGTCTTCTGGCATGTTGACTCGTCTGCTCTGCGAGTGTATGACCGGGTTGTG 4895
Qy 1502 CTTGGTGTAGCTTACGCGCCCGGAGACCAAGTGTAGCTTACGAGTGTATGACCGGGTTGTG 1561
Db 4896 CTTGGTGTAGCTTACGCGCCCGGAGACTACAGTGTAGCTTACGAGTGTATGACCGGGTTGTG 4955
Qy 1562 CGGACTTCCGCTGTGCCAAGACCATCTTGAATTTTGGAGGGGCTCTTACGGGCTCTCA 1621
Db 4956 CGGACTTCCGCTGTGCCAAGACCATCTTGAATTTTGGAGGGGCTCTTACAGGCGCTCA 5015
Qy 1622 CCCATAGAGCGCCACTTCTTATCCCGAGACCAAGAGAGTGGGAAACCTTCCCTATC 1681
Db 5016 CTCATATAGATGCGCCACTTCTTATCCCGAGACCAAGAGAGTGGGAAACCTTCCCTATC 5075
Qy 1682 TGGTAGGCTTACCAAGCACCGTGTGCGTGTAGAGTCAAGCCCTCCCGCGTGGGAGCC 1741
Db 5076 TGGTAGGCTTACCAAGCACCGTGTGCGTGTAGAGTCAAGCCCTCCCGCGTGGGAGCC 5135
Qy 1742 AGATGTGGAAGTGTGTATCGCTCTCAAGCCACCTCCATGGGCGCAACACTCTGCTAT 1801
Db 5136 AGATGTGGAAGTGTGTATCGCTCTCAAGCCACCTCCATGGGCGCAACACTCTGCTAT 5195
Qy 1802 ATAGACTGGGCGCTGTCCAGATGAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCA 1861
Db 5196 ACAGACTGGGCGCTGTTCAGATGAAATCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCA 5255
Qy 1862 TGACATGTATGTGCGCTGAGCTGAGGTGCTGAGTCAAGTCAAGTCAAGTCAAGTCAAGTCA 1921
Db 5256 TGACATGTATGTGCGGCGACTGAGGTGCTGAGTCAAGTCAAGTCAAGTCAAGTCAAGTCA 5315
Qy 1922 TTCTGGTGTCTTTGGCGCGCTATTGCTCTATCCAGAGTGTGCTGTATAGTAGGTAGGA 1981
Db 5316 TTCTGGTGTCTTTGGCGCGCTATTGCTCTATCCAGAGTGTGCTGTATAGTAGGTAGGA 5375
Qy 1982 TTGTCTTGTTCGGAAGCGCGCAATCATACCGGACAGGGAAGTCTCTTACCGGAGTTGCG 2041
Db 5376 TCGTCTTGTTCGGAAGCGCGCAATCATACCGGAGAGTCTCTTACCGGAGTTGCG 5435
Qy 2042 ATGAATGGAAGAGTGTCT 2059
Db 5436 ATGAGATGGAAGAGTGTCT 5453
```

RESULT 14

AAT12710

ID AAT12710 standard; cDNA; 9401 BP.

XX

AC AAT12710;

XX DT 25-MAR-2003 (revised)  
DT 15-MAY-1996 (first entry)  
XX DE Hepatitis C virus polyprotein.  
XX KW Non-A non-B hepatitis virus; NANBHV; HCV; antigen; detection; diagnosis;  
XX OS antibodies; ds.  
XX OS Hepatitis C virus.  
XX FH Key Location/Qualifiers  
FT CDS 342..9378  
FT /\*tag= a  
XX EP693687-A1.  
XX PN 24-JAN-1996.  
XX PD 03-APR-1991; 95EP-00114016.  
XX PF 04-APR-1990; 90US-00504352.  
XX PR (CHIR ) CHIRON CORP.  
XX PA Houghton M, Choo Q, Kuo G;  
XX PI WPI; 1996-117956/13.  
XX PS P-PSDB; AAR90931.  
XX PT Combinations of synthetic Hepatitis C Virus antigens - provide more  
XX effective diagnosis of Non-A, Non-B Hepatitis.  
XX PS Disclosure; Fig 1(A-Y); 53pp; English.  
XX CC The combination comprises an HCV antigen from the C domain (pref. C22 -  
CC AAR90936) and at least one HCV antigen from the NS3 (pref. C33c -  
CC AAR90932), NS4 (pref. C100 - AAR90933), S (pref. S2 - AAR90935) or NS5  
CC (AAR90934) domain. The antigens may in the form of a fusion protein, a  
CC simple physical mixture, or the individual antigens commonly bound to a  
CC solid matrix. They are pref. prepd. by recombinant DNA techniques  
CC (primers are given in AAT12711-T12716), but can be synthesised or  
CC isolated from HCV using affinity chromatography. (Updated on 25-MAR-2003  
CC to correct PF field.)  
XX SQ Sequence 9401 BP; 1883 A; 2860 C; 2673 G; 1985 T; 0 U; 0 Other;

Query Match 86.7%; Score 1786; DB 2; Length 9401;  
Best Local Similarity 91.7%; Pred. No. 0;  
Matches 1886; Conservative 0; Mismatches 170; Indels 0; Gaps 0;

Qy 2 TGGCGCCTATCAGCGCCTATGCCAGCAGACAGAGGGGCGCTTTGGGATGCATATCACCA 61  
Db 3418 TGGCGCCTATCAGCGCCTATGCCAGCAGACAGAGGGGCGCTTTGGGATGCATATCACCA 3477

Qy 62 GCTTACCGCGCGGACAAACACAGGTGAGGGTGCAGATCGTGTCAACTGCTG 121  
Db 3478 GCCTAATCGCGGACAAACACAGGTGAGGGTGCAGATCGTGTCAACTGCTG 3537

Qy 122 CCCAGACTTTCTTGGCAACCTGCATTAACCGGGGTGTGTGGACTGTCTACCATGGAGCCG 181  
Db 3538 CCCAAACCTTCTTGGCAACCTGCATTAACCGGGGTGTGTGGACTGTCTACCATGGAGCCG 3597

Qy 182 GAACAGAGCATTTGCTACCTACCTAAGGTCTGTGTATCCAGATGTACCAATGTGAGCC 241  
Db 3598 GAACAGAGCATTTGCTACCTACCTAAGGTCTGTGTATCCAGATGTATACCAATGTAGACC 3657

Qy 242 AAGACCTCGTAGGCTGGCCCGCTCCCAAGGTGCCCTCATTAACACCATGCACTTGGCG 301  
Db 3658 AAGACCTTGTGGGCTGGCCCGCTCCCAAGGTGCCCTCATTAACACCATGCACTTGGCG 3717

Qy 302 GCTCCTCGGACCTTTACCTGCTCAACGAGGACGCGCATGTCTTCTGTGCGCGGAGCGG 361  
|||||







Qy	1502	CTTGGTATAGCTTACGCCCGCGAGACCAACAGTTAGGTTACGAGCATACATGAACACC	1561
Db	4918		
Qy	1562	CGGGCTTCCCGTGTGCCAAGACCATCTTGAATTTGGGAGGGCGTCTTTACGGGTCTCA	1621
Db	4978		
Qy	1622	CCACATAGACGCCCACTTCTCTATCCAGACAAAGCAGAGTGGGGAACCTTTCCCTATC	1681
Db	5038		
Qy	1682	TGGTAGCGTACCAAGCCACCGTGTGGCTAGAGCTCAAGCCCTCCCGCTCGTGGGACC	1741
Db	5098		
Qy	1742	AGATGTGGAAGTGTTCATCCGTCCTCAAGCCCACTTCCATGGGCGCAACACCTCTGCTAT	1801
Db	5158		
Qy	1802	ATAGACTGGGCGCTGTCCAGAATGAAGTCACTGACGACCCAGTCACCAAGTATATCA	1861
Db	5218		
Qy	1862	TGACATGTATGTGGGCTGACCTGGAGTGTCTACGAGTACCTGGGTGCTCTGTTGGCGCG	1921
Db	5278		
Qy	1922	TTCTGGCTGTTTGGCGCGGTATTTGCCCTATCCACAGGCTCGGTGGTTCATAGTAGGTAGGA	1981
Db	5338		
Qy	1982	TTGTCTTGTCCGGAAGCGGCAATCATACCCGACAGGGAAGTCTCTACCGGGAGTTGG	2041
Db	5398		
Qy	2042	ATGAAATGGAAGAGTGCT	2059
Db	5458		

Search completed: February 27, 2005, 18:30:06  
Job time : 1064 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 27, 2005, 17:06:13 ; Search time 370 Seconds  
(without alignments)  
9114.504 Million cell updates/sec

Title: US-09-930-591-1

Perfect score: 2061

Sequence: 1 atggcgctatcacggcta.....atgaatggaagagtgtga 2061

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued\_Patents\_NA.\*  
1: /cgn2\_6/ptodata/1/ina/5A\_COMB.seq.\*  
2: /cgn2\_6/ptodata/1/ina/5B\_COMB.seq.\*  
3: /cgn2\_6/ptodata/1/ina/6A\_COMB.seq.\*  
4: /cgn2\_6/ptodata/1/ina/6B\_COMB.seq.\*  
5: /cgn2\_6/ptodata/1/ina/PTUS\_COMB.seq.\*  
6: /cgn2\_6/ptodata/1/ina/backfiles1.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1786	86.7	7310	3	US-08-444-818-74
2	1786	86.7	9379	3	US-09-388-874-1
3	1786	86.7	9379	4	US-09-916-359-1
4	1786	86.7	9401	1	US-07-910-760-9
5	1786	86.7	9401	1	US-08-440-519-9
6	1786	86.7	9401	3	US-08-440-549-9
7	1786	86.7	9401	3	US-08-823-895A-25
8	1785.2	86.6	9401	5	PCT-US91-02225-9
9	1784.4	86.6	6785	3	US-08-444-818-65
10	1784.4	86.6	8316	3	US-08-444-818-88
11	1784.4	86.6	8987	3	US-08-444-818-137
12	1784.4	86.6	9185	3	US-08-444-818-122
13	1784.4	86.6	9185	3	US-08-444-818-123
14	1784.4	86.6	9379	3	US-08-444-818-176
15	1781.2	86.4	2058	4	US-09-881-239-2
16	1781.2	86.4	5360	3	US-08-444-818-53
17	1779.6	86.3	2058	4	US-09-881-654-1
18	1779.6	86.3	2058	4	US-10-637-323-1
19	1774.8	86.1	12980	3	US-08-811-566-5
20	1774.8	86.1	12980	3	US-09-034-756-5
21	1773.2	86.0	9646	3	US-08-811-566-1
22	1773.2	86.0	9646	3	US-09-034-756-1
23	1765.2	85.6	9599	3	US-09-014-416-2
24	1765.2	85.6	9599	3	US-09-014-416-6
25	1762	85.5	9416	3	US-08-823-895A-26
26	1762	85.5	9416	4	US-10-104-966-13
27	1750.8	84.9	9401	2	US-08-432-693-1
					Sequence 74, Appl
					Sequence 1, Appl
					Sequence 1, Appl
					Sequence 9, Appl
					Sequence 9, Appl
					Sequence 9, Appl
					Sequence 25, Appl
					Sequence 9, Appl
					Sequence 88, Appl
					Sequence 137, App
					Sequence 122, App
					Sequence 123, App
					Sequence 176, App
					Sequence 2, Appl
					Sequence 53, Appl
					Sequence 1, Appl
					Sequence 5, Appl
					Sequence 1, Appl
					Sequence 1, Appl
					Sequence 2, Appl
					Sequence 6, Appl
					Sequence 26, Appl
					Sequence 13, Appl
					Sequence 1, Appl

Sequence 19, Appl  
Sequence 19, Appl  
Sequence 7, Appl  
Sequence 69, Appl  
Sequence 69, Appl  
Sequence 69, Appl  
Sequence 69, Appl  
Sequence 85, Appl  
Sequence 85, Appl  
Sequence 1, Appl  
Sequence 3, Appl  
Sequence 6, Appl  
Sequence 10, Appl  
Sequence 7, Appl  
Sequence 22, Appl  
Sequence 7, Appl

## ALIGNMENTS

RESULT 1  
US-08-444-818-74  
; Sequence 74, Application US/08444818  
; Patent No. 6150087  
; GENERAL INFORMATION:  
; APPLICANT: Chien, David Y.  
; APPLICANT: Rutter, William J.  
; TITLE OF INVENTION: NANBV Diagnostics and Vaccines  
; NUMBER OF SEQUENCES: 777  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Chiron Corporation  
; STREET: 4560 Horton Street  
; CITY: Emeryville  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 94608-2916  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/444,818  
; FILING DATE:  
; CLASSIFICATION: 424  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/08/403,590  
; FILING DATE: 14-MAR-1995  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Harbin, Alisa A.  
; REGISTRATION NUMBER: 33,895  
; REFERENCE/DOCKET NUMBER: 0110.002  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (508)359-3876  
; TELEFAX: (508)359-3885  
; INFORMATION FOR SEQ ID NO: 74:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 7310 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: cDNA  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: 3..7310  
US-08-444-818-74

Query Match 86.7%; Score 1786; DB 3; Length 7310;  
Best Local Similarity 91.7%; Pred. No. 0;  
Matches 1888; Conservative 0; Mismatches 170; Indels 0; Gaps 0;

QY 2 TGCGGCTATACGGCTATGCCCCAGACAGAGAGGGGCTTTTGGGATGCATAATCACCA 61  
Db 1729 TGGCGCCCATACGGGTAGCGCCAGCAGACAGAGGGGCTTCTAGGGTCATTAATCACCA 1788  
QY 62 GCTTGACCGGCGGGCAAAACACAGGTGGAGGTGAGGTTCAGATCGTGTCAACTGCTG 121  
Db 1789 GCCTAACTGGCGGGCAAAACCAAGTGGAGGTGAGGTCCAGATTGTGCAACTGCTG 1848  
QY 122 CCAGACTTTCTTGGCAACCTGATTAACCGGGGTGTGGAGTGTCTACCATGGAGCGG 181  
Db 1849 CCACAACTTTCTTGGCAACCTGATTAACCGGGGTGTGGAGTGTCTACCATGGAGCGG 1908  
QY 182 GAACAAAGGACATTTGGTCACTTAAGGCTCTCTTATCCAGATGTACACCAATGTGGACC 241  
Db 1909 GAACGAGGACCATCGGTCACTTAAGGCTCTCTTATCCAGATGTATCCAAATGTAGACC 1968  
QY 242 AAGACCTCTGCTAGGCTGGCGCTCCCAAGGTGCGCGCTCAATTAACCAATGCACTTGGC 301  
Db 1969 AAGACCTTTGGCTGGCGCTCCCAAGGTAGCGCTCAATTAACCAATGCACTTGGC 2028  
QY 302 GCTCTCGGACCTTTACCTGGTCAAGAGGACCGCGATGTCTCTGTCGGCGGAGCGG 361  
Db 2029 GCTCTCGGACCTTTACCTGGTCAAGAGGACCGCGATGTCTCTGTCGGCGGAGCGG 2088  
QY 362 GTGATGCGAGGGGAGCGCTGCTGTCGGCGCGGCGCAATTTCTACTTTGMAAGGCTCTCGG 421  
Db 2089 GTGATGCGAGGGGAGCGCTGCTGTCGGCGCGGCGCAATTTCTACTTTGMAAGGCTCTCGG 2148  
QY 422 GAGGCGCTCTGCTGTCGGCGGAGGACATGCGGTAGGCAATTTCAAGCGCGGATGCA 481  
Db 2149 GGGGTCTGCTGTTGTCGGCGGAGGACGCGGTGGGCAATTTAGGGCGCGGTGTGCA 2208  
QY 482 CCGGTGAGGTGGCTAAGCGGTGGACTTCAATCCCGTAGAGACTTTAGAGCAACCATGA 541  
Db 2209 CCGGTGAGGTGGCTAAGCGGTGGACTTCAATCCCGTAGAGACTTTAGAGCAACCATGA 2268  
QY 542 GGTCCCGGCTGTTCTCAGACAACTCTCCCGACAGCAGTGGCCCGAGAGTACCAAGTGG 601  
Db 2269 GGTCCCGGCTGTTCTCAGACAACTCTCCCGACAGTGGCCCGAGAGTTCAGAGTGG 2328  
QY 602 CCACCTGATGCTCCCGAGCGGAGGTAAGAGCACCAAGGTCCCGGCGCGATACGAG 661  
Db 2329 CTCACTTCCATGCTCCCGAGCGGAGGCAAAAGCAACCAAGGTCCCGGCTGCATATGCA 2388  
QY 662 CTAGGGCTACAGGTGCTGCTGCTCAACCCCTCCGTTGCTGCAACAAATGGGCTTTGGTG 721  
Db 2389 CTAGGGCTATAGGTGCTAGTACTCAACCCCTCTGTTGCTGCAACAACTGGGCTTTGGTG 2448  
QY 722 CTTTACATGTCAGAGGCCATGGGATTTGATCCTAACATCAGGACTGGGGTGAGGACAAATTA 781  
Db 2449 CTTTACATGTCAGAGGCTCATGGGATCGATCCTAACATCAGGAGCGGGGTGAGAACAAATTA 2508  
QY 782 CTACTGCGACCCGATCAGATATTCACCTACGCGCAAGTTCTTTGCGGAGCGGGGTGTT 841  
Db 2509 CCAGTGGCAGCCCATCAGCTACTCCACTACGCAAGTTCTTTGCGGAGCGGGGTGCT 2568  
QY 842 CAGGGGTGCTTATGACATTAATTTGTGACGAGTGCCTCACTCCAGGATGCAACATCCA 901  
Db 2569 CAGGGGTGCTTATGACATTAATTTGTGACGAGTGCCTCACTCCAGGATGCAACATCCA 2628  
QY 902 TCTTGGGCATTTGGACATGCTCTTTGACCAAGCAGAGACCGGGGGGAGAGTCACTGACTGTC 961  
Db 2629 TCTTGGGCATTTGGACATGCTCTTTGACCAAGCAGAGACTGCGGGGGGAGAGTCTGTTGTC 2688  
QY 962 TCGCCACCGGTACCCCTCCGGGCTCGGTCACTGTGCCCCATCTTAACATCAGAGAGTTG 1021  
Db 2689 TCGCCACCGGTACCCCTCCGGGCTCGGTCACTGTGCCCCATCTTAACATCAGAGAGTTG 2748  
QY 1022 CTCTGTCCACTACCGGAGAGATCCCTTTTATGGCAGGCTATTCCTCTTGAAGCAATTA 1081  
Db 2749 CTCTGTCCACTACCGGAGAGATCCCTTTTATGGCAGGCTATTCCTCTTGAAGCAATTA 2808

QY 1082 AGGGGGGAGACATCTCATCTTCTGCCACTCAAGAAGAGTGCAGAGCTGCGCGCAA 1141  
Db 2809 AGGGGGGAGACATCTCATCTTCTGCCACTCAAGAAGAGTGCAGAGCTGCGCGCAA 2868  
QY 1142 AACTGGTCCGTTGGCGTCAATGCGGTGCTTACTACGCGGCTTGTGATGTGCTGCTCA 1201  
Db 2869 AGCTGGTCCATTTGGGCATCAATGCGGTGCTTACTACGCGGCTTGTGATGTGCTGCTCA 2928  
QY 1202 TCCGACCAAGTGTGAGCTGCTGCTGGCAACTGACGCCCTCATGACCGGCTTTACCG 1261  
Db 2929 TCCGACCAAGTGTGAGCTGCTGCTGGCAACTGACGCCCTCATGACCGGCTTTACCG 2988  
QY 1262 GCGACTTCGATTTGGGTAGATGCTGCAACACGCTGTGTACCCAGACAGTTCAGGC 1321  
Db 2989 GCGACTTCGACTCGGTGATAGACTGCAATACGCTGTGTACCCAGACAGTTCAGGC 3048  
QY 1322 TTGACCTTACCTTACCATTTGAGACATCAAGTTCCTCCAGGATGCTGCTCCCGTACTC 1381  
Db 3049 TTGACCTTACCTTACCATTTGAGACATCAAGTTCCTCCAGGATGCTGCTCCCGTACTC 3108  
QY 1382 AACGTCCGGGTAGGACTGCGAGAGGGAAGCCAGGCACTTACAGATTTGTGGCACCGGGG 1441  
Db 3109 AACGTCCGGGTAGGACTGCGAGAGGGAAGCCAGGCACTTACAGATTTGTGGCACCGGGG 3168  
QY 1442 AGCGTCTTCTGCGCATGTTTGAATGCTGCTCTGCGAGTGTATGACCGGGTGTG 1501  
Db 3169 AGCGTCTTCTGCGCATGTTTGAATGCTGCTCTGCGAGTGTATGACCGAGGCTGTG 3228  
QY 1502 CTTGTTATGAGCTTAGCGCGCGGAGACACAGTTAGGCTAGAGCATACATGAACACCC 1561  
Db 3229 CTTGTTATGAGCTTAGCGCGCGGAGACACAGTTAGGCTAGAGCATACATGAACACCC 3288  
QY 1562 CGGGACTTCCCGTGTGCCAAGACCATCTTTGAATTTTGGGAGGCGCTTTTACGGGCTCA 1621  
Db 3289 CGGGGCTTCCCGTGTGCCAAGACCATCTTTGAATTTTGGGAGGCGCTTTTACAGGCTCA 3348  
QY 1622 CCCACATAGACGCCCATCTTCTATCCAGACAAAGCAGAGTGGGGAAAACTTTCCCTATC 1681  
Db 3349 CTTATATAGATGCCCATCTTCTATCCAGACAAAGCAGAGTGGGGAGAACCTTTCTTACC 3408  
QY 1682 TGGTAGCTTACCAAGCCAGGCTGCGGTAGAGCTCAAGCCCTCCCGTCTGTTGGGAGC 1741  
Db 3409 TGGTAGCTTACCAAGCCAGGCTGCGGTAGAGCTCAAGCCCTCCCGTCTGTTGGGAGC 3468  
QY 1742 AGATGTGAAGTGTCTTGATCCGCTCTCAAGCCACCCCTCCATGGGCCAAACACCTCTGCTAT 1801  
Db 3469 AGATGTGAAGTGTCTTGATTCGCTCTCAAGCCACCCCTCCATGGGCCAAACACCTCTGCTAT 3528  
QY 1802 ATAGACTGGGCGCTGTCAGAAATGAAGTCACTCTGACGACCCAGTCAACCAATATATCA 1861  
Db 3529 ACAGACTGGGCGCTGTTCAAGAAATGAAGTCACTCTGACGACCCAGTCAACCAATATATCA 3588  
QY 1862 TGCATGATGTGCGCTGACCTGGAGTGTGCTGACAGTACCTGGGTGCTGTTGGGCGG 1921  
Db 3589 TGCATGATGTGCGCTGACCTGGAGTGTGCTGACAGTACCTGGGTGCTGTTGGGCGG 3648  
QY 1922 TTCTGCTGCTTTTGGCGGCTATTGCTTATCCAGGCTGCTGCTGCTATAGTAGTAGGA 1981  
Db 3649 TCTGCTGCTTTTGGCGGCTATTGCTTATCCAGGCTGCTGCTGCTATAGTAGTAGGA 3708  
QY 1982 TTGCTTTGTCGGAAAGCGCGCAATCATACCCGACAGGAGTCTCTTACCGGAGTTG 2041  
Db 3709 TCGTCTTGTTCGGGAAAGCGCGCAATCATACCTGACAGGAGTCTCTTACCGAGAGTTG 3768  
QY 2042 ATGAATGGAAGTGTCT 2059  
Db 3769 ATGAGATGGAAGTGTCT 3786

RESULT 2

US-09-388-874-1

; Sequence 1, Application US/09388874

; Patent No. 6284249

## GENERAL INFORMATION:

APPLICANT: Veronique Barban  
TITLE OF INVENTION: VACCINE COMPOSITION FOR PREVENTING OR  
TREATING C HEPATITIS  
FILE REFERENCE: PMCF97-03A  
CURRENT APPLICATION NUMBER: US/09/388,874  
CURRENT FILING DATE: 1999-09-02  
EARLIER APPLICATION NUMBER: PCT/FR98/00448  
EARLIER FILING DATE: 1998-03-06  
EARLIER APPLICATION NUMBER: 97/02,887  
EARLIER FILING DATE: 1997-03-06  
NUMBER OF SEQ ID NOS: 2  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 1  
LENGTH: 9379  
TYPE: DNA  
ORGANISM: Virus  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (320)... (9352)  
US-09-388-874-1

Query Match 86.7%; Score 1786; DB 3; Length 9379;

Best Local Similarity 91.7%; Pred. No. 0;  
Matches 1888; Conservative 0; Mismatches 170; Indels 0; Gaps 0;

Qy	2	TGGCGCCTATCACGCGCTATGCCAGCAGACAAGGGGCTTTTGGGATGCAATACCA	61
Db	3396	TGGCGCCCATCACGCGGTACGCCAGCAGACAAGGGGCTCTCTAGGGTGCATATACCA	3455
Qy	62	GCTTGACGGCGGGACAAAACAGGTGGAGGTGAGGTTTCAGATCGTGTCAACTGCTG	121
Db	3456	GCCTTAACCTGGCGGGACAAAACAGGTGGAGGTGAGGTTTCAGATCGTGTCAACTGCTG	3515
Qy	122	CCAGACTTCTTGGCAACCTGCAATTAACGGGTGTTGGACCTGTACCATCGAGCGG	181
Db	3516	CCCAACCTTCTGGCAACCTGCAATTAACGGGTGTTGGACCTGTACCATCGAGCGG	3575
Qy	182	GAAACAAGGACCAATTCGCTACCTAAGGCTCTGTTATCCAGATGATACCAATGTGACC	241
Db	3576	GAACAGGACCAATTCGCTACCTAAGGCTCTGTTATCCAGATGATACCAATGTGACC	3635
Qy	242	AAGACCTGTAGGCTGGCGGCTCCCAAGGTGCGCTCATTAACACCAATGCACTTGGG	301
Db	3636	AAGACCTTGTGGCTGGCGGCTCCCAAGGTGCGCTCATTAACACCAATGCACTTGGG	3695
Qy	302	GCTCCTCGGACCTTACCTGTGACAGGACGCGGATGTCATTCCTGTGGCGCGAGGG	361
Db	3696	GCTCCTCGGACCTTACCTGTGACAGGACGCGGATGTCATTCCTGTGGCGCGAGGG	3755
Qy	362	GTGATGGCAGGGGAGGCTGCTTTTCGCGCGCGGCTATCTTCTTGAAGGCTCCTCGG	421
Db	3756	GTGATGACGGGGAGGCTGCTGTCGCGCGCGGCTATCTTCTTGAAGGCTCCTCGG	3815
Qy	422	GAGGCCCTCTGCTGTGGCGCGAGGATGCGGTAGGCAATTCAGAGCGCGGTATGCA	481
Db	3816	GGGGTCCGCTGTTGTGGCGCGGAGCGCGGTGGGCATATTTAGGGCGCGGTGTGCA	3875
Qy	482	CCCGTGGAGTGGCTTAAGCGGTGACCTTATCCCGGTAGAGAGCTTAGAGACCAACATGA	541
Db	3876	CCCGTGGAGTGGCTTAAGCGGTGACCTTATCCCGGTGGAGAACTTAGAGACCAACATGA	3935
Qy	542	GGTCCCGGCTTCTCAGACAACCTCTCCCAACAGAGAGTCCCGCAGAGCTTACCAAGTGG	601
Db	3936	GGTCCCGGCTTCTCAGAGTAACCTCTCTCCACAGTAGTGCCTCCAGAGCTTCCAGGTGG	3995
Qy	602	CCCACTGTGATGTCTCCACCGGACGGTAAAGACCAAGGTCCCGCGCGCATACGAG	661
Db	3996	CTCACTCCATGTCTCCACAGGACGGGCAAAAGCAACCAAGGTCCCGCGCTGATATGAG	4055
Qy	662	CTCAGGCTACAAGGTGCTGTCTCAACCCCTCGGTGCTGCAACATGGGCTTTGGTG	721
Db	4056	CTCAGGCTACAAGGTGCTGTCTCAACCCCTCGGTGCTGCAACATGGGCTTTGGTG	4115

Qy	722	CTTACATGTCCAAGGCCCATGGGATTTGATCTTAACATCAGAGTGGGTGAGGCAATTA	781
Db	4116	CTTACATGTCCAAGGCTCATGGGATCGATCTTAACATCAGAGCGGGTGAGAACATTA	4175
Qy	782	CTACTGCGAGCGCGATCAGTATTTCCACCTACCGCAAGTTCTTTGCCAGCGCGGTGT	841
Db	4176	CCACTGGCAGCGCCCATCAGTACTCCACCTACCGCAAGTTCTTTGCCAGCGCGGTGT	4235
Qy	842	CAGGGGTGCTTATGACATAATATTTGTGACGAGTCCCACTCCACGATGCAACATCCA	901
Db	4236	CGGGGGCGCTTATGACATAATATTTGTGACGAGTCCCACTCCACGATGCAACATCCA	4295
Qy	902	TCTTGGCATTTGGCATCTCTTGAACCAAGCAGAGACCGCGGGGCGAGACTGACTGTGC	961
Db	4296	TCTTGGCATCTGGCATCTCTTGAACCAAGCAGAGACTGCGGGGCGAGACTGCTGTGTGC	4355
Qy	962	TCGCCACCGCTACCCCTCCGGGCTCGTCACTGTGCCCATCTTAACATCAGAGAGTTG	1021
Db	4356	TCGCCACCGCACCCCTCCGGGCTCGTCACTGTGCCCATCTTAACATCAGAGAGTTG	4415
Qy	1022	CTCTGTCCACTACCGGAGAGATCCCTTTTATGGCAAGGCTATTCCCTTTGAAGCAATTA	1081
Db	4416	CTCTGTCCAACCGGAGAGATCCCTTTTATGGCAAGGCTATTCCCTTCGAAGTATCA	4475
Qy	1082	AGGGGGGAGACATCTCATCTTCTGCACTCAAAAGAAAGTGCAGAGCTCCCGCA	1141
Db	4476	AGGGGGGAGACATCTCATCTTCTGCACTCAAAAGAAAGTGCAGAGCTCCCGCA	4535
Qy	1142	AACGTGTCGGTGGGGTCAATGCCGTGCTTACTACCGGGCTTTGATGTGCTGCTCA	1201
Db	4536	AGCTGTGCGATTTGGGCATCAATGCCGTGCTTACTACCGGGCTTTGATGTGCTGCTCA	4595
Qy	1202	TCCGACACGAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG	1261
Db	4596	TCCGACACGAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG	4655
Qy	1262	GCGACTTCGACTCGGTGATAGACTGCAACACGCTGTGTCCACGACAGTGCAGCTTCA	1321
Db	4656	GCGACTTCGACTCGGTGATAGACTGCAACACGCTGTGTCCACGACAGTGCAGCTTCA	4715
Qy	1322	TTGACCTTACCTTACCAATTTGAGCAATCAGCTTCCCGAGGATGCTGTCTCCCGTACTC	1381
Db	4716	TTGACCTTACCTTACCAATTTGAGCAATCAGCTTCCCGAGGATGCTGTCTCCCGTACTC	4775
Qy	1382	AAGTGGGGGTAGGACTGGCAGAGGGAAGCAGCATCTACAGATTTGTGSCACCGGGG	1441
Db	4776	AAGTGGGGGTAGGACTGGCAGAGGGAAGCAGCATCTACAGATTTGTGSCACCGGGG	4835
Qy	1442	AGGTCCTCTTGGCATGTTTGAATCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1501
Db	4836	AGGTCCTCTTGGCATGTTTGAATCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	4895
Qy	1502	CTTGTGATGAGCTTACGCGCGCGAGACCAAGTGTAGGCTACGAGCATACATGAACACC	1561
Db	4896	CTTGTGATGAGCTTACGCGCGCGAGACCAAGTGTAGGCTACGAGCATACATGAACACC	4955
Qy	1562	CGGACTTCCCGTGTGCAAGACCATCTGCAATTTTGGGAGGGGCTCTTTACGGGTCTCA	1621
Db	4956	CGGACTTCCCGTGTGCAAGACCATCTGCAATTTTGGGAGGGGCTCTTTACGGGTCTCA	5015
Qy	1622	CCACATAGAGCGCCACTTCTTATCCAGACAAGCAGAGTGGGGAACCTTCCCTATC	1681
Db	5016	CTCATATAGATGCCACTTCTTATCCAGACAAGCAGAGTGGGGAACCTTCTTACC	5075
Qy	1682	TGGTAGGCTACCAAGCCACCGTGTGCGTACAGCTCAAGCCCTCCCGCTGCTGGGACC	1741
Db	5076	TGGTAGGCTACCAAGCCACCGTGTGCGTACAGCTCAAGCCCTCCCGCTGCTGGGACC	5135
Qy	1742	AGATGTGGAAGTGTGATCCGCTCAAGCCCACTCCATGGGCCCAACCTCTGCTAT	1801
Db	5136	AGATGTGGAAGTGTGATCCGCTCAAGCCCACTCCATGGGCCCAACCTCTGCTAT	5195





## RESULT 4

[illegible]



```
/
/ OTHER INFORMATION: can also be Thr."
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: 4167
/ OTHER INFORMATION: /note= "This amino acid position
/ OTHER INFORMATION: can also be Leu."
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: 4323
/ OTHER INFORMATION: /note= "This amino acid position
/ OTHER INFORMATION: can also be Val."
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: 4701
/ OTHER INFORMATION: /note= "This amino acid position
/ OTHER INFORMATION: can also be Tyr."
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: 4752
/ OTHER INFORMATION: /note= "This amino acid position
/ OTHER INFORMATION: can also be Ser."
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: 5970
/ OTHER INFORMATION: /note= "This amino acid position
/ OTHER INFORMATION: can also be Gly."
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: 6183
/ OTHER INFORMATION: /note= "This amino acid position
/ OTHER INFORMATION: can also be His."
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: 6186
/ OTHER INFORMATION: /note= "This amino acid position
/ OTHER INFORMATION: can also be Cys."
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: 6402
/ OTHER INFORMATION: /note= "This amino acid position
/ OTHER INFORMATION: can also be Val."
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: 7386
/ OTHER INFORMATION: /note= "This amino acid position
/ OTHER INFORMATION: can also be Ser."
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: 7494
/ OTHER INFORMATION: /note= "This amino acid position
/ OTHER INFORMATION: can also be Phe."
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: 7497
/ OTHER INFORMATION: /note= "This amino acid position
/ OTHER INFORMATION: can also be Ala."
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: 7845
/ OTHER INFORMATION: /note= "This amino acid position
/ OTHER INFORMATION: can also be Phe."
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: 8409
/ OTHER INFORMATION: /note= "This amino acid position
/ OTHER INFORMATION: can also be Gly."
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: 9102
/ OTHER INFORMATION: /note= "This amino acid position
/ OTHER INFORMATION: can also be Gly."
/ FEATURE:
/ NAME/KEY: misc_feature
```

```
/
/ LOCATION: 9327
/ OTHER INFORMATION: /note= "This amino acid position
/ OTHER INFORMATION: can also be Pro."
/ US-08-440-519-9

Query Match      86.7%; Score 1786; DB 1; Length 9401;
Best Local Similarity 91.7%; Pred. No. 0;
Matches 1888; Conservative 0; Mismatches 170; Indels 0; Gaps 0;

QY      2  TGGCGCCTATACGGCCTATGCCAGCAGACAGAGGGCGCTTTTGGGATGCATATACCA 61
DB      3418 TGGCGCCCATCAGCGGTACGCCACAGACACAGGGGCGCTCTAGGGTGCAATATACCA 3477

QY      62  GCTTGACCGCGCGGACAAAACAGGTGGAGGGTGAGGTTCAAGTCGTGTCAACTGCTG 121
DB      3478 GCCTAACTGGCGGGACAAAACCAAGTGGAGGGTGAGGTCAGATGTTGTCAACTGCTG 3537

QY      122 CCCAGACTTTTGTGGCAACTGCAATTAACGGGGTGTGTTGGAGTGTCTACCATGGAGCGG 181
DB      3538 CCCAAACCTTCTGTGGCAACGTGCATCAATGGGGTGTGCTGGAGTGTCTACACGGGCGG 3597

QY      182 GAACAGGACCATTGGCTCACCTAAGGTCCTGTTATCCAGATGTACACCAATGTGGACC 241
DB      3598 GAACAGGACCATTCGGCTCACCAAGGGTCTGTGCATCCAGATGTATACCAATGTAGACC 3657

QY      242 AAGACCTCGTAGGCTGGCGCGCTCCCAAGGTGCCCGCTCATTAAACAACCATGCACTTGGC 301
DB      3658 AAGACCTTGTGGCTGGCGCGCTCCGCAAGTAGCGCTCATTTGACACCTGCACTTGGC 3717

QY      302 GCTCTCGAACCTTTACTCTGGTCAAGAGCAGCCGATGTCATCTCTGTGCGCGCAGGG 361
DB      3718 GCTCTCGAACCTTTACTTGGTCACGAGGACGCGCATGTCATCTCTGCGCGCGGG 3777

QY      362 GTGATCGCAGGGGCGAGCTGCTTTCCCGCGGCTATCTTACTTTGAAGGCTCTCCGG 421
DB      3778 GTGATAGCAGGGGCGAGCTGCTGTCCCGCGGCGCAATTCCTACTTTGAAAGGCTCTCTCGG 3837

QY      422 GAGGCGCTCTGCTGTGCCCGCAGGACATGCGCTAGGCATATTTCAGAGCGCGGCTATGCA 481
DB      3838 GGGGTCCGCTGTGTGCCCGCGGGGACGCGCTGGGCATATTATTAGGCGCGCGGTGTGCA 3897

QY      482 CCGGTGAGTGGCTAAGCGGGTGGACTTTCATCCCGTATAGAGCTTAGAGAACCAATGA 541
DB      3898 CCGGTGAGTGGCTAAGCGGGTGGACTTTATCTCTGTGGAGAACCTAGAGAACCAATGA 3957

QY      542 GGTCCCCGGGTGTTCTCAGACAACTCTCCCGCAGCAGTGGCCCCAGAGCTACCAAGTG 601
DB      3958 GGTCCCCGGGTGTTCCAGGATACTCTCTCCACAGTAGTGCCCCAGAGCTTCAGGTGG 4017

QY      602 CCACCTGCATGCTCCACCGCAGCGGTAAAGACCAAGGTCCCGCGCGCATAGCAG 661
DB      4018 CTCACCTCCATGCTCCACAGGCGGGCNAAGCACCAGGTCCCGGCTGCATATGCAG 4077

QY      662 CTCAGGGCTACAAGGTGCTGGTCTCAACCCCTCGTGTGCTGCAACAANGGGCTTTGGTG 721
DB      4078 CTCAGGGCTATAAGGTGCTAGTACTCAACCCCTCTGTGCTGCAACAACCTGGGCTTTGGTG 4137

QY      722 CTTACATGTCCAAGGCCCATGGGATTCCTTAACATCAGGACTGGGCTGAGSACAATTA 781
DB      4138 CTTACATGTCCAAGGGCTCATGGGATCGATCTTAACATCAGGACCGGGGTGAGAACAAATTA 4197

QY      782 CTACTGGCAGCCGATCAGTATTTCCACTACGGCAAGTTCCTTTGCCAGCGGGGTGTT 841
DB      4198 CCACTGGCAGCCCATCACTACTCCACTACGGCAAGTTCCTTTGCCAGCGGGGTGCT 4257

QY      842 CAGGGGGTCTTATGACATAATAATTTGTGACAGTGCCACTCCACGGATGCAACATCCA 901
DB      4258 CGGGGGGCGCTTATGACATAATAATTTGTGACAGTGCCACTCCACGGATGCCACATCCA 4317

QY      902 TCTTGGGCATTGGCACTGCTCTTGACCAAGCAGAGACCGCGGGGCGAGACTCACTGTGC 961
DB      4318 TCTTGGGCATCGGCACCTGTCTTTGACCAAGCAGAGACTGCGGGGGCGAGACTGTTGTGC 4377
```

QY 962 TCGCCACCGCTACCCCTCGGGCTCGCTCACTGTGCCCCCATCTTAACATCGAGAGGTTG 1021  
Db 4378 TCGCCACCGCCACCCCTCGGGCTCGCTCACTGTGCCCCCATCTTAACATCGAGAGGTTG 4437  
QY 1022 CTCTGTCCATACCGGAGAGATCCCTTTTATGGCAAGCTATTCCTCTTGAAGCAATTA 1081  
Db 4438 CTCTGTCCATACCGGAGAGATCCCTTTTATGGCAAGCTATTCCTCTTGAAGCAATTA 4497  
QY 1082 AGGGGGGAGAGATCTCATCTCTTCCCACTCAAGAAGAAGTGGCAGAGCTCGCGCAA 1141  
Db 4498 AGGGGGGAGAGATCTCATCTCTTCCCACTCAAGAAGAAGTGGCAGAGCTCGCGCAA 4557  
QY 1142 AACTGTGCGCTGGCGCTCAATCGCTGTCTACTACCGGGCTTGTATGTCTCCGTCA 1201  
Db 4558 AGCTGTGCGATTTGGGCTCAATCGCTGTCTACTACCGGGCTTGTATGTCTCCGTCA 4617  
QY 1202 TCCGACCAAGTGTGCGCTGTGCGTGGCAACTGACGCTCATGACCGCTTTTACCG 1261  
Db 4618 TCCGACCAAGTGTGCGCTGTGCGTGGCAACTGACGCTCATGACCGCTTTTACCG 4677  
QY 1262 CGGACTTCGATTCGGTGTATAGACTGCAACAGTGTGTCAACCGAGACAGTTCAGCC 1321  
Db 4678 CGGACTTCGATTCGGTGTATAGACTGCAACAGTGTGTCAACCGAGACAGTTCAGCC 4737  
QY 1322 TTGACCTTACCTTCACTTACGAGCAATCAAGCTTCCCGAGATGCTGTCTCCGTACTC 1381  
Db 4738 TTGACCTTACCTTCACTTACGAGCAATCAAGCTTCCCGAGATGCTGTCTCCGTACTC 4797  
QY 1382 AACGTCGGGTAGGACTGCGAGAGGAGGAGGAGGATCTACAGATTTTGGCAACCGGGG 1441  
Db 4798 AACGTCGGGTAGGACTGCGAGAGGAGGAGGAGGATCTACAGATTTTGGCAACCGGGG 4857  
QY 1442 AGCGTCTTTCGCACTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1501  
Db 4858 AGCGTCTTTCGCACTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 4917  
QY 1502 CTTGTATGAGTTACGCGCGGAGACCAAGTTAGGCTACGAGATACATGAAACACC 1561  
Db 4918 CTTGTATGAGTTACGCGCGGAGACCAAGTTAGGCTACGAGATACATGAAACACC 4977  
QY 1562 CGGAGCTTCCGCTGTCAGAGACCATCTTGAATTTGGAGGGCTCTTTACGGGTCTCA 1621  
Db 4978 CGGAGCTTCCGCTGTCAGAGACCATCTTGAATTTGGAGGGCTCTTTACGGGTCTCA 5037  
QY 1622 CCCATAGAGCCGCTTCTTATCCCAAGAAAGAGAGTGGGAAACCTTCCCTATC 1681  
Db 5038 CTCATATAGATGCCACTTCTATCCCAAGAAAGAGAGTGGGAAACCTTCCCTATC 5097  
QY 1682 TGGTAGCGTACCAAGCCACCGTGTGCGTCAAGCTCAAGCCCTCCCGCTCGTGGGACC 1741  
Db 5098 TGGTAGCGTACCAAGCCACCGTGTGCGTCAAGCTCAAGCCCTCCCGCTCGTGGGACC 5157  
QY 1742 AGATGTGAAGTGTGATCGCTCTCAAGCCACCCCTCCATGGGCGCAACCTCTGCTAT 1801  
Db 5158 AGATGTGAAGTGTGATCGCTCTCAAGCCACCCCTCCATGGGCGCAACCTCTGCTAT 5217  
QY 1802 ATAGACTGGGCGCTGTCCAGAAATGAAGTCAACCTGACGACCCAGTCAACCAAGTATCA 1861  
Db 5218 ACAGACTGGGCGCTGTCCAGAAATGAAGTCAACCTGACGACCCAGTCAACCAAGTATCA 5277  
QY 1862 TGACATGATGTGCGCTGACTGAGGTCGTCAAGAGTACTGAGTGTCTGCTGGCGCG 1921  
Db 5278 TGACATGATGTGCGCTGACTGAGGTCGTCAAGAGTACTGAGTGTCTGCTGGCGCG 5337  
QY 1922 TTCTGGCTGCTTGGCGCGTATTCCTTATCCAGGCTGCTGCTGCTAGTAGTAGGA 1981  
Db 5338 TTCTGGCTGCTTGGCGCGTATTCCTTATCCAGGCTGCTGCTGCTAGTAGTAGGA 5397  
QY 1982 TTGTCTTCTCGGAAAGCGGCAATCATACCGAGGGAAGTCTCTTACCGGAGTTCG 2041  
Db 5398 TCGTCTTCTCGGAAAGCGGCAATCATACCGAGGGAAGTCTCTTACCGGAGTTCG 5457  
QY 2042 ATGAATGGAAGAGTGCT 2059

Db 5458 ATGAGATGGAAGAGTGCT 5475

## RESULT 6

US-08-440-549-9  
; Sequence 9, Application US/08440549  
; Patent No. 6312889  
; GENERAL INFORMATION:  
; APPLICANT: Houghton, Michael  
; APPLICANT: Choo, Qui-Lim  
; APPLICANT: Kuo, George  
; TITLE OF INVENTION: Combinations of Hepatitis C virus (HCV)  
; TITLE OF INVENTION: Antigens for Use in Immunoassays for Anti-HCV Antibodies  
; NUMBER OF SEQUENCES: 12  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Chiron Corporation  
; STREET: P.O. Box 8097 (Int. Prop. R-440)  
; CITY: Emeryville  
; STATE: CA  
; COUNTRY: U.S.A.  
; ZIP: 94662-8097  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/440,549  
; FILING DATE: 12-MAY-1995  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/910,760  
; FILING DATE: 07-JUL-1992  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Blackburn Esq., Robert P.  
; REGISTRATION NUMBER: 30,447  
; REFERENCE/DOCKET NUMBER: 0101.002  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (510) 601-2702  
; TELEFAX: (510) 655-3542  
; INFORMATION FOR SEQ ID NO: 9:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 9401 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: linear  
; MOLECULE TYPE: DNA (genomic)  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: 342..9374  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: 366  
; OTHER INFORMATION: /note= "This amino acid position  
; OTHER INFORMATION: can also be Arg."  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: 372  
; OTHER INFORMATION: /note= "This amino acid position  
; OTHER INFORMATION: can also be Thr."  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: 867  
; OTHER INFORMATION: /note= "This amino acid position  
; OTHER INFORMATION: can also be Thr."  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: 1341  
; OTHER INFORMATION: /note= "This amino acid position  
; OTHER INFORMATION: can also be Val."  
; FEATURE:  
; NAME/KEY: misc\_feature

/ LOCATION: 2148 /note= "This amino acid position  
/ OTHER INFORMATION: can also be Ile."  
/ FEATURE:  
/ NAME/KEY: misc\_feature  
/ LOCATION: 2883  
/ OTHER INFORMATION: /note= "This amino acid position  
/ OTHER INFORMATION: can also be Asn."  
/ FEATURE:  
/ NAME/KEY: misc\_feature  
/ LOCATION: 3681  
/ OTHER INFORMATION: /note= "This amino acid position  
/ OTHER INFORMATION: can also be Ser."  
/ FEATURE:  
/ NAME/KEY: misc\_feature  
/ LOCATION: 3690  
/ OTHER INFORMATION: /note= "This amino acid position  
/ OTHER INFORMATION: can also be Thr."  
/ FEATURE:  
/ NAME/KEY: misc\_feature  
/ LOCATION: 4167  
/ OTHER INFORMATION: /note= "This amino acid position  
/ OTHER INFORMATION: can also be Leu."  
/ FEATURE:  
/ NAME/KEY: misc\_feature  
/ LOCATION: 4323  
/ OTHER INFORMATION: /note= "This amino acid position  
/ OTHER INFORMATION: can also be Val."  
/ FEATURE:  
/ NAME/KEY: misc\_feature  
/ LOCATION: 4701  
/ OTHER INFORMATION: /note= "This amino acid position  
/ OTHER INFORMATION: can also be Tyr."  
/ FEATURE:  
/ NAME/KEY: misc\_feature  
/ LOCATION: 4752  
/ OTHER INFORMATION: /note= "This amino acid position  
/ OTHER INFORMATION: can also be Ser."  
/ FEATURE:  
/ NAME/KEY: misc\_feature  
/ LOCATION: 5970  
/ OTHER INFORMATION: /note= "This amino acid position  
/ OTHER INFORMATION: can also be Gly."  
/ FEATURE:  
/ NAME/KEY: misc\_feature  
/ LOCATION: 6183  
/ OTHER INFORMATION: /note= "This amino acid position  
/ OTHER INFORMATION: can also be His."  
/ FEATURE:  
/ NAME/KEY: misc\_feature  
/ LOCATION: 6186  
/ OTHER INFORMATION: /note= "This amino acid position  
/ OTHER INFORMATION: can also be Cys."  
/ FEATURE:  
/ NAME/KEY: misc\_feature  
/ LOCATION: 6402  
/ OTHER INFORMATION: /note= "This amino acid position  
/ OTHER INFORMATION: can also be Val."  
/ FEATURE:  
/ NAME/KEY: misc\_feature  
/ LOCATION: 7386  
/ OTHER INFORMATION: /note= "This amino acid position  
/ OTHER INFORMATION: can also be Ser."  
/ FEATURE:  
/ NAME/KEY: misc\_feature  
/ LOCATION: 7494  
/ OTHER INFORMATION: /note= "This amino acid position  
/ OTHER INFORMATION: can also be Phe."  
/ FEATURE:  
/ NAME/KEY: misc\_feature  
/ LOCATION: 7497  
/ OTHER INFORMATION: /note= "This amino acid position  
/ OTHER INFORMATION: can also be Ala."

/ FEATURE:  
/ NAME/KEY: misc\_feature  
/ LOCATION: 7845  
/ OTHER INFORMATION: /note= "This amino acid position  
/ OTHER INFORMATION: can also be Phe."  
/ FEATURE:  
/ NAME/KEY: misc\_feature  
/ LOCATION: 8409  
/ OTHER INFORMATION: /note= "This amino acid position  
/ OTHER INFORMATION: can also be Gly."  
/ FEATURE:  
/ NAME/KEY: misc\_feature  
/ LOCATION: 9102  
/ OTHER INFORMATION: /note= "This amino acid position  
/ OTHER INFORMATION: can also be Gly."  
/ FEATURE:  
/ NAME/KEY: misc\_feature  
/ LOCATION: 9327  
/ OTHER INFORMATION: /note= "This amino acid position  
/ OTHER INFORMATION: can also be Pro."  
/ US-08-440-549-9  
Query Match 86.7%; Score 1786; DB 3; Length 9401;  
Best Local Similarity 91.7%; Pred. No. 0;  
Matches 1888; Conservative 0; Mismatches 170; Indels 0; Gaps 0;  
QY 2 TGGCGCTATCAGCGCTATGCCAGCAGACAGGGGCTTTTGGGATGCAATACCA 61  
DB 3418 TGGCGCCCATCAGCGGTACGCCAGCAGACAGGGGCTTCTAGGTTGATATACCA 3477  
QY 62 GCTTGACCGCGCGGACAAAACAGGTGGAGGTTCAGATCGTGTCAACTGCTG 121  
DB 3478 GCCTAATGCGCGGACAAAACAGTGGAGGTGAGTCCAGATTGTCAACTGCTG 3537  
QY 122 CCAGACTTTCTTGGCAACTGCAITTAACGGGGTGTGGACTGTCTACCATGGAGCGG 181  
DB 3538 CCCAAACCTTCTCTGGCAAGTGCATCAATGGGGTGTGCTGGACTGTCTACACGGGCGG 3597  
QY 182 GAACAGGACCATTTGGTGCACCTTAAGGGTCTCTTATCCAGATGTACACCAATGTGACC 241  
DB 3598 GAACAGGACCATTCGGCTCACCAAGGTCTCTTATCCAGATGTATACCAATGTAGACC 3657  
QY 242 AAGACCTCTAGGCTGGCGCGCTCCCAAGGTGCGCGCTCATTAACCAATGCATTGCG 301  
DB 3658 AAGACCTTGTGGCTGGCGCGCTCCGCAAGGTAGCGGCTCATTAACCAATGCATTGCG 3717  
QY 302 GCTCTCGGACCTTTACCTGGTGCACAGGACCGCGATGTCTTCTGTGCGCGGCGG 361  
DB 3718 GCTCTCGGACCTTTTACCTGGTGCACAGGACCGCGATGTCTTCTGTGCGCGGCGG 3777  
QY 362 GTGATGCGAGGGGCGAGCTGCTTTGCGCGCGCTATCTTACTTTGAAGGCTCCTCGG 421  
DB 3778 GTGATGCGAGGGGCGAGCTGCTGCTGCGCGCGCGCTATCTTACTTTGAAGGCTCCTCGG 3837  
QY 422 GAGGCGCTCTGCTGTGCTGCGCGGAGCATGCGCTAGGCATATTTCAGAGCGCGGTATGCA 481  
DB 3838 GGGGTGCGCTGTTGTGCGCGCGGCGACGCGGTGGGCATATTTCAGGCGCGGTGTGCA 3897  
QY 482 CCGTGGAGTGGCTAAGGGGGTGGACTTCATCCCGGTAGAGAGCTTAGAGACCAACATGA 541  
DB 3898 CCGTGGAGTGGCTAAGGGGGTGGACTTTATCCCTGTGGAGAACCTTAGAGACCAACATGA 3957  
QY 542 GGTCCCGGCTTCTCAGACAACTCTCCCGCAGCAGTCCCGCAGAGCTACCAAGTGG 601  
DB 3958 GGTCCCGGCTTCTCAGAGTAACCTCTCTCCACAGTAGTGGCCCGAGAGCTTCAGGTGG 4017  
QY 602 CCACCTGCATGCTCCCGCAGCGGTAAAGACCAAGGTCCCGCGCGCATACGACAG 661  
DB 4018 CTCACCTCCATGCTCCCGCAGCGCGGCAAAAGCACAAGGTCCCGCGGTGCATATGACAG 4077  
QY 662 CTAGGGCTAAGAGGTGCTGGTCTCAACCCCTCGGTGCTGCTGCAACATGGGGCTTTGGTG 721  
DB 4078 CTAGGGCTAAGAGGTGCTAGTACTCAACCCCTCTGTGCTGCAACATGGGGCTTTGGTG 4137



QY 722 CTTACATGTCAGAGCCCATCGGATTTGATCTTAACATCAGAGCTGGGGTGAGGCAATTA 781  
Db 4138 CTTACATGTCAGAGCCCATCGGATTCGATCTTAACATCAGAGCCGGGGTGAGGCAATTA 4197  
QY 782 CTACTGGCAGCCCGATCAGCTATTCCACTACGGCAAGTTCTTGGCCGACGGCGGGTGT 841  
Db 4198 CCACTGGCAGCCCGATCAGCTATTCCACTACGGCAAGTTCTTGGCCGACGGCGGGTGT 4257  
QY 842 CAGGGGTGCTTATGACATAATAATTTGTGACGAGTGCCACTCCACGAGATGCAACATCCA 901  
Db 4258 CGGGGGCGCTTATGACATAATAATTTGTGACGAGTGCCACTCCACGAGATGCCACATCCA 4317  
QY 902 TCTTGGGATTTGGCACTGTCTTGAACAGCAGAGACCGCGGGGGCGAGACTGACTGTGC 961  
Db 4318 TCTTGGGATTTGGCACTGTCTTGAACAGCAGAGACTGCGGGGGCGAGACTGCTGTGTC 4377  
QY 962 TCGCCACCGCTACCCCTCCGGGCTCCGTCAGTGTGCCCCCTTAACATCGAGGAGGTG 1021  
Db 4378 TCGCCACCGCCCTCCGGGCTCCGTCAGTGTGCCCCCTTAACATCGAGGAGGTG 4437  
QY 1022 CTCTGTCCACTACCGGAGAGATCCCTTTTATGCAAGGCTATTCCCTTGAACAATTA 1081  
Db 4438 CTCTGTCCACACCGGAGAGATCCCTTTTACGCAAGGCTATCCCTCGAAGTATCA 4497  
QY 1082 AGGGGGGAGACATCTCATCTTCTTGCCACTCAAAAGAAAGTGCAGAGCTGCCGCCAA 1141  
Db 4498 AGGGGGGAGACATCTCATCTTCTTGCTATTCAAAAGAAAGTGCAGAGCTGCCGCCAA 4557  
QY 1142 AACTGTCGGCTTGGCGCTCAATCCGCTGCTTACTACCGGGCTTGTGATGTGCCGTC 1201  
Db 4558 AGCTGGTCGATTTGGGATCAATCCGCTGCTTACTACCGGGCTTGTGATGTGCCGTC 4617  
QY 1202 TCCGACACGAGTGTGCTGCTGCTGGCAACTGACGCGCTCATGACCGGCTTTTACCG 1261  
Db 4618 TCCGACACGAGGATGTGCTGCTGCTGGCAACCGATGCCCTCATGACCGGCTTATACG 4677  
QY 1262 GCGATTCGATTTGGTGATGACTGCAACACGCTGTGTCAACCAGACAGTGCAGCTTACGCC 1321  
Db 4678 GCGATTCGACTCGGTGATGACTGCAATACGCTGTGTCAACCAGACAGTGCAGCTTACGCC 4737  
QY 1322 TTGACCTTACCTTACATTTGAGACATACGCTTCCGAGGATGCTGTCTCCGCTACTC 1381  
Db 4738 TTGACCTTACCTTACATTTGAGACATACGCTTCCGAGGATGCTGTCTCCGCTACTC 4797  
QY 1382 AACGTGGGGTAGGACTGGCAGAGGGAAGCAGGATCTACAGATTTGTGCAACGGGG 1441  
Db 4798 AACGTGGGGAGGACTGGCAGAGGGAAGCAGGATCTACAGATTTGTGCAACGGGG 4857  
QY 1442 AGCGTCTCTTGGCATGTTTGAATCTGCTGTCTCTGCGAGTGTATGACGCGGGTTGTG 1501  
Db 4858 AGCGCCCTCCGGCATGTTGAGTCTGCTCGTCTCTGAGTGTATGACGAGGCTGTG 4917  
QY 1502 CTTGGTATGAGTTACCGCCCGCAGACACAGTTAGGCTACGAGCATATACATGAACACC 1561  
Db 4918 CTTGGTATGAGTTACCGCCCGCAGACACAGTTAGGCTACGAGCATATACATGAACACC 4977  
QY 1562 CGGAGCTTCCGCTGTCAGAACCATCTGAAATTTTGGGAGGGCTTTTACGGGCTCA 1621  
Db 4978 CGGGGCTTCCGCTGTCAGAACCATCTGAAATTTTGGGAGGGCTTTTACGGGCTCA 5037  
QY 1622 CCCACATAGACGCCACTTCTATCCAGACAAAGCAGAGTGGGGAACCTTCCCTTATC 1681  
Db 5038 CTATATAGATGCCACTTCTATCCAGACAAAGCAGAGTGGGGAACCTTCCCTTACC 5097  
QY 1682 TGGTAGGATACCAAGCCACCGTGTGCGCTAGAGCTCAAGGCCCTTCCCGCTGTGGGACC 1741  
Db 5098 TGGTAGGATACCAAGCCACCGTGTGCGCTAGAGCTCAAGGCCCTTCCCGCTGTGGGACC 5157  
QY 1742 AGATGTGGAGTGTCTTCACTCGCTTCAAGCCACCTTCCATGGCCCAACCTCTGCTAT 1801  
Db 5158 AGATGTGGAGTGTCTTCACTCGCTTCAAGCCACCTTCCATGGCCCAACCTCTGCTAT 5217

QY 1802 ATAGACTGGGGCTGTCCAGAAATGAAGTCACTGACGACCCAGTCAACCAAGTATATCA 1861  
Db 5218 ACAGACTGGGGCTGTTCAGAAATGAAGTCACTGACGACCCAGTCAACCAATATCA 5277  
QY 1862 TGACATGTATGTGGCTGACTGGAGTCTGTCAGAGTACCTGGTGTCTGTTGGCGGG 1921  
Db 5278 TGACATGTATGTGGCGACTGGAGTCTGTCAGAGACCTGGTGTCTGTTGGCGGG 5337  
QY 1922 TTCTGGCTGTCTTGGCGCGCTATTGCTTATCCATCCAGAGTGGTGTCTATAGTAGGA 1981  
Db 5338 TCCTGGCTGTCTTGGCGCGCTATTGCTTATCCATCCAGAGTGGTGTCTATAGTAGGA 5397  
QY 1982 TTGTCTTGTCCGAAAGCCGCGCAATCATACCCGACAGGAGTCTCTTACCGGAGTTTCG 2041  
Db 5398 TCCTTGTCTGGGAAGCCGCGCAATCATACCTGACAGGAGTCTCTTACCGAGAGTTTCG 5457  
QY 2042 ATGAAATGGAAGAGTGCT 2059  
Db 5458 ATGAGATGGAAGAGTGCT 5475

## RESULT 7

US-08-823-895A-25

; Sequence 25, Application US/08823895A

; Patent No. 6433159

; GENERAL INFORMATION:

; APPLICANT: Kevin P. Anderson

; TITLE OF INVENTION: Compositions And Methods For

; TITLE OF INVENTION: Treatment Of Hepatitis C Virus-Associated Diseases

; NUMBER OF SEQUENCES: 27

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Jane Massey Licata, Esq.

; STREET: 66 E. Main Street

; CITY: Marlton

; STATE: NJ

; COUNTRY: USA

; ZIP: 08053

; COMPUTER READABLE FORM:

; MEDIUM TYPE: DISKETTE, 3.5 INCH, 1.44 Mb STORAGE

; COMPUTER: IBM 486

; OPERATING SYSTEM: WINDOWS FOR WORKGROUPS

; SOFTWARE: WORDPERFECT 5.1

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/823,895A

; FILING DATE: March 17, 1997

; CLASSIFICATION: 514

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 08/453,085

; FILING DATE: May 30, 1995

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 07/945,289

; FILING DATE: September 10, 1992

; ATTORNEY/AGENT INFORMATION:

; NAME: Jane Massey Licata

; REGISTRATION NUMBER: 32,257

; REFERENCE/DOCKET NUMBER: ISPH-0203

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (609) 779-2400

; TELEFAX: (609) 810-1454

; INFORMATION FOR SEQ ID NO: 25:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 9401

; TYPE: Nucleic

; STRANDEDNESS: Single

; TOPOLOGY: Linear

; ANTI-SENSE: NO

; US-08-823-895A-25

Query Match 86.7%; Score 1786; DB 3; Length 9401;

Best Local Similarity 91.7%; Pred. No. 0;

Matches 1889; Conservative 0; Mismatches 170; Indels 0; Gaps 0;

QY 2 TGGCGCTATCAGCGGCTATGCCAGTATGCCAGCAGCAAGGGGCGCTTTTGGGATGCATATACCA 61



APPLICANT: KUO, GEORGE  
TITLE OF INVENTION: COMBINATIONS OF HEPATITIS C VIRUS  
TITLE OF INVENTION: ANTIBODIES  
TITLE OF INVENTION: ANTIBODIES  
NUMBER OF SEQUENCES: 10  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Morrison & Foerster  
STREET: 545 Middlefield Road, Suite 200  
CITY: Menlo Park  
STATE: CA  
COUNTRY: USA  
ZIP: 94025  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent-In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US91/02225  
FILING DATE: 19910329  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: CIOTTI, THOMAS E.  
REGISTRATION NUMBER: 21,013  
REFERENCE/DOCKET NUMBER: 2300-0101.44  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 327-7250  
TELEFAX: (415) 327-2951  
TELEX: 706141  
INFORMATION FOR SEQ ID NO: 9:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 9401 base pairs  
TYPE: NUCLEIC ACID  
STRANDEDNESS: unknown  
TOPOLOGY: unknown  
MOLECULE TYPE: DNA (genomic)  
PCT-US91-02225-9

Query Match 86.6%; Score 1785.2; DB 5; Length 9401;  
Best Local Similarity 91.5%; Pred. No. 0;  
Matches 1883; Conservative 6; Mismatches 169; Indels 0; Gaps 0;

Qy	2	TGGGCGCTATCAGCGCTATGCCAGAGACCAAGGGCGCTTTGGGATGCATATACCA	61
Db	3418	TGGCGCCCATCAGCGCTATGCCAGAGACCAAGGGCGCTTTGGGATGCATATACCA	3477
Qy	62	GCTTACCGCGCGGACAAACAGGTGGAGGTGAGTTTCAGATGCTGCAACTGCTG	121
Db	3478	GCCTAACTGGCGCGGACAAACAGGTGGAGGTGAGTTTCAGATGCTGCAACTGCTG	3537
Qy	122	CCAGACTTTCTTGGCAACCTGCATTAACGGGGTGTGTGGACTGTCTACCATGGAGCG	181
Db	3538	CCCAAACTTCTTGGCAACCTGCATTAACGGGGTGTGTGGACTGTCTACCATGGAGCG	3597
Qy	182	GAAACAGGACATTTGGCTACCTAAGGCTCTGTATCCAGATGACCAACATGTGACC	241
Db	3598	GAAACAGGACATTTGGCTACCTAAGGCTCTGTATCCAGATGACCAACATGTGACC	3657
Qy	242	AAGACCTCGTAGGCTGGCGCGCTCCCAAGGTGGCGCTCATTAACCATGCACTTGG	301
Db	3658	AAGACCTCGTAGGCTGGCGCGCTCCCAAGGTGGCGCTCATTAACCATGCACTTGG	3717
Qy	302	GCTCCTCGGACTTTTACCTGCTACAGGACAGCGCGATGTCATTCCTGTGGCGCGGG	361
Db	3718	GCTCCTCGGACTTTTACCTGCTACAGGACAGCGCGATGTCATTCCTGTGGCGCGGG	3777
Qy	362	GTGATGGCAGGGGAGCGCTGCTTTTCGCCCGCGCTATCTTACTTGAAGGCTCTCGG	421
Db	3778	GTGATGGCAGGGGAGCGCTGCTTTTCGCCCGCGCTATCTTACTTGAAGGCTCTCGG	3837
Qy	422	GAGGCGCTCTGCTGTGCCCGCAGACATGCCGTAGGCATATTCAGAGCGCGGTATGCA	481
Db	3838	GGGGTCCGCTGTGTGTGCCCGCAGACATGCCGTAGGCATATTCAGAGCGCGGTATGCA	3897

Qy	482	CCCGTGAGTGGCTAAGCGCGGTGCACTTTCATCCCGGTAGAGAGCTTAGAGACAAACCATGA	541
Db	3898	CCCGTGAGTGGCTAAGCGCGGTGCACTTTCATCCCGGTAGAGAGCTTAGAGACAAACCATGA	3957
Qy	542	GGTCCCGGTGTTCTCAGACAACTCTCCACACAGAGTGGCCAGAGCTTACCAAGTGG	601
Db	3958	GGTCCCGGTGTTCTCAGACAACTCTCTCCACAGTAGTGGCCAGAGCTTCCAGGTGG	4017
Qy	602	CCACCTGCATGCTCCACCGGACAGCGGTAAAGACACCAAGGTCCCGGCCCATACGAG	661
Db	4018	CTACCTCCATGCTCCACAGGACAGCGGCAAAAGCAACAGGTCCCGGCTGCATATGAG	4077
Qy	662	CTCAGGGCTACAAGTGTCTGCTCAACCCCTCCGTGTGTGCAACAAATGGCTTTGGTG	721
Db	4078	CTCAGGGCTACAAGTGTCTGCTCAACCCCTCTGTGTGTGCAACAACTGGCTTTGGTG	4137
Qy	722	CTTACATGTCNAAGGCCCATGGGATGATCTTAACATCAGAGCTGGGGTGAAGCAATTA	781
Db	4138	CTTACATGTCNAAGGCCCATGGGATGATCTTAACATCAGAGCTGGGGTGAAGCAATTA	4197
Qy	782	CTACTGGCAGCCGATCACGTTATTCACCTACGGCAAGTTCCTTGGCCGACGGCGGTGT	841
Db	4198	CCACTGGCAGCCGATCACGTTATTCACCTACGGCAAGTTCCTTGGCCGACGGCGGTGT	4257
Qy	842	CAGGGGTGCTTATGACATAAATTTGTGACGAGTGCCTCCACCGATGCAACATCCA	901
Db	4258	CGGGGGCGCTTATGACATAAATTTGTGACGAGTGCCTCCACCGATGCAACATCCA	4317
Qy	902	TCCTGGGCATTTGGCACTGCTTTCACCAAGCAGAGACCGGGGGGCGAGACTGACTGTC	961
Db	4318	TCCTGGGCATTTGGCACTGCTTTCACCAAGCAGAGACCGGGGGGCGAGACTGACTGTC	4377
Qy	962	TCGCCACCGCTACCCCTCCGGGCTCCGTCTACTGTGCCCTCTCTAAATCAGAGAGTTG	1021
Db	4378	TCGCCACCGCTACCCCTCCGGGCTCCGTCTACTGTGCCCTCTCTAAATCAGAGAGTTG	4437
Qy	1022	CTCTGTCCACTACCGGAGAGATCCCTTTTATGCAAGGCTATTCCTCTTTGAAGCAATTA	1081
Db	4438	CTCTGTCCACTACCGGAGAGATCCCTTTTATGCAAGGCTATTCCTCTTTGAAGCAATTA	4497
Qy	1082	AGGGGGGAGACATCTCATCTTCTGCCACTCAAAAGAGAGTGGCAGCTCCCGCAA	1141
Db	4498	AGGGGGGAGACATCTCATCTTCTGCCACTCAAAAGAGAGTGGCAGCTCCCGCAA	4557
Qy	1142	AACGTGTCGGTGGGGCTCAATGCGGTGCTTACTACCGGGCTTGTATGTGCTCCGTCA	1201
Db	4558	AGCTGTCGGATTTGGGCATCAATGCCGTGGCTACTACCGGGCTTGTATGTGCTCCGTCA	4617
Qy	1202	TCCGACCCAGTGGTGTGCTGCTGGCAACTGACGGCTCTCATGACCGGCTTTACCG	1261
Db	4618	TCCGACCCAGTGGTGTGCTGCTGGCAACTGACGGCTCTCATGACCGGCTTTACCG	4677
Qy	1262	GGACTTCGATTCGGTGTATGACTGCAACAGTGTGTCAACAGAGAGTGCAGCTTCAGCC	1321
Db	4678	GGACTTCGATTCGGTGTATGACTGCAACAGTGTGTCAACAGAGAGTGCAGCTTCAGCC	4737
Qy	1322	TTGACCCCTACTTACCATTTGAGACATCAAGTTCCTCCAGGATGCTGTCTCCGCTATC	1381
Db	4738	TTGACCCCTACTTACCATTTGAGACATCAAGTTCCTCCAGGATGCTGTCTCCGCTATC	4797
Qy	1382	AAAGTCGGGGTAGGACTGGCAGAGGGAAGCGAGCATCTACAGATTTGTGGCACCGGGGG	1441
Db	4798	AAAGTCGGGGTAGGACTGGCAGAGGGAAGCGAGCATCTACAGATTTGTGGCACCGGGGG	4857
Qy	1442	AGGTCCTCTTCTGGCATGTTTGAATCGTCTCTCTCGGAGTGTATGACGCGGGTGTG	1501
Db	4858	AGGTCCTCTTCTGGCATGTTTGAATCGTCTCTCTCGGAGTGTATGACGCGGGTGTG	4917
Qy	1502	CTTGGTATGAGCTTACCGCCCGGAGACCAAGTTAGCTACGAGCATATCATGAACACC	1561
Db	4918	CTTGGTATGAGCTTACCGCCCGGAGACCAAGTTAGCTACGAGCATATCATGAACACC	4977

QY 1562 CGGACTCCCGTGTGCCAAGACCATCTGAATTTGGAGGGCGCTTTACGGGTCTCA 1621  
Db CCGGGCTCCCGTGTGCCAGGACCATCTGAATTTGGAGGGCGCTTTACAGGCTCA 5037  
QY 1622 CCACATAGAGCCCACTTCTATCCAGACAAAGCAGAGTGGGAAAACTTCCCTATC 1681  
Db CTCATATAGATGCCCACTTCTATCCAGACAAAGCAGAGTGGGAGAACCTTCTTACC 5097  
QY 1682 TGGTAGGTACCAAGCCACCGTGTGCCGTAGAGCTCAAGCCCTCCCGCTCGTGGGACC 1741  
Db TGGTAGGTACCAAGCCACCGTGTGCCGTAGAGCTCAAGCCCTCCCGCTCGTGGGACC 5157  
QY 1742 AGATGTGGAAGTGTTCGATCCGCTCAAGCCCACTTCCATGGGCCCAACCACTCTGCTAT 1801  
Db AGATGTGGAAGTGTTCGATCCGCTCAAGCCCACTTCCATGGGCCCAACCACTCTGCTAT 5217  
QY 1802 ATAGACTGGGGCTGTCCAGATGAAGTCAACCTGACGCAACCAAGTCAACCAATATCA 1861  
Db ACAGACTGGGGCTGTTCAGATGAAGTCAACCTGACGCAACCAAGTCAACCAATATCA 5277  
QY 1862 TGACATGTATGTGGCTGACTGAGGTCTGACAGTACCTGGGTGCTGTTGGGGCG 1921  
Db TGACATGTATGTGGCTGACTGAGGTCTGACAGTACCTGGGTGCTGTTGGGGCG 5337  
QY 1922 TTCTGGCTGTGTTGGCGCGTATTTCCCTATCCAGGCTGCGTGGTCAATAGTAGTAGGA 1981  
Db TCCTGGCTGTGTTGGCGCGTATTTCCCTATCCAGGCTGCGTGGTCAATAGTAGTAGGA 5397  
QY 1982 TTCTGTGTCGGAAGCCGGCAATCATACCGACAGGGAAGTCTCTACCGGGAGTTGG 2041  
Db TCCTGTGTCGGAAGCCGGCAATCATACCGACAGGGAAGTCTCTACCGGGAGTTGG 5457  
QY 2042 ATGAATGGAAGAGTGTCT 2059  
Db ATGAGATGGAAGAGTGTCT 5475

## RESULT 9

US-08-444-818-65  
; Sequence 65, Application US/08444818  
; Patent No. 6150087  
; GENERAL INFORMATION:  
; APPLICANT: Chien, David Y.  
; APPLICANT: Rutter, William J.  
; TITLE OF INVENTION: NANBV Diagnostics and Vaccines  
; NUMBER OF SEQUENCES: 777  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Chiron Corporation  
; STREET: 4560 Horton Street  
; CITY: Emeryville  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 94608-2916  
; COMPUTER READABLE FORM: disk  
; MEDIUM TYPE: Floppy  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/444,818  
; FILING DATE:  
; CLASSIFICATION: 424  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/08/403,590  
; FILING DATE: 14-MAR-1995  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Harbin, Alisa A.  
; REGISTRATION NUMBER: 33,895  
; REFERENCE/DOCKET NUMBER: 0110.002  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (508)359-3876  
; TELEFAX: (508)359-3885  
; INFORMATION FOR SEQ ID NO: 65:

; SEQUENCE CHARACTERISTICS:  
; LENGTH: 6785 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: cDNA  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: 3..6785  
; US-08-444-818-65  
  
Query Match 86.6%; Score 1784.4; DB 3; Length 6785;  
Best Local Similarity 91.7%; Pred. No. 0;  
Matches 1887; Conservative 0; Mismatches 171; Indels 0; Gaps 0;  
  
QY 2 TGGCGCTATACAGGGCTATGCCCAGACAGCAAGGGGCCCTTTGGATGCAATACCA 61  
Db 1204 TGGCGCCATACAGGGCTATGCCCAGACAGCAAGGGGCCCTTTGGATGCAATACCA 1263  
QY 62 GCTTGACCGCGGAGCAAAAACCAAGGTGAGGGTTCAGATCGTGTCAACTGCTG 121  
Db 1264 GCTTAACCTGCGCGGAGCAAAAACCAAGGTGAGGGTTCAGATCGTGTCAACTGCTG 1323  
QY 122 CCAGACTTTCTTGGCAACCTTGCAATTAACGGGGTGTGTTGGACTGTCTACCATGGAGCGG 181  
Db 1324 CCAGAACTTCTTGGCAACCTTGCAATTAACGGGGTGTGTTGGACTGTCTACCATGGAGCGG 1383  
QY 182 GAACAGGACCAATGGGTCACTAAGGTCTCTGTTATCCAGATGTACCAATGTGGACC 241  
Db 1384 GAACAGGACCAATGGGTCACTAAGGTCTCTGTTATCCAGATGTATACCAATGTAGACC 1443  
QY 242 AAGACCTCTAGCTGGCGCGCTCCCAAGGTGCGCGCTATTAAACACCATGCACCTTGG 301  
Db 1444 AAGACCTCTAGCTGGCGCGCTCCCAAGGTGCGCGCTATTAAACACCATGCACCTTGG 1503  
QY 302 GCTCTCTGGAACCTTTACCTGGTCAAGGACCGCCGATGCTATTCCTGTGCCCGACGGG 361  
Db 1504 GCTCTCTGGAACCTTTACCTGGTCAAGGACCGCCGATGCTATTCCTGTGCCCGACGGG 1563  
QY 362 GTGATGGCAGGGGAGCGCTTTCGCGCGCGCTATCTCTTACTTGAAGGCTCTCTCG 421  
Db 1564 GTGATGGCAGGGGAGCGCTTTCGCGCGCGCTATCTCTTACTTGAAGGCTCTCTCG 1623  
QY 422 GAGGCGCTCTGCTGTCGCCCGCAGACATGCGTAGGCATATTCAGAGCGCGGTATGA 481  
Db 1624 GGGGTCCGCTGTGTCGCCCGCAGACCGCGTGGGCATATTTAGGGCCGCGGTGTGA 1683  
QY 482 CCCGTGGAGTGGCTAAGCGGTGGACTTTCATCCCGCTAGAGAGCTTAGAGACCAACATGA 541  
Db 1684 CCCGTGGAGTGGCTAAGCGGTGGACTTTCATCCCGCTAGAGAGCTTAGAGACCAACATGA 1743  
QY 542 GGTCCCGGTGTTCTCAGACAACTCTCTCCACAGCAGTGCCTCCAGAGCTACCAAGTGG 601  
Db 1744 GGTCCCGGTGTTCTCAGACAACTCTCTCCACAGCAGTGCCTCCAGAGCTACCAAGTGG 1803  
QY 602 CCCACTGCATGCTCCCAACCGCAGCGGTAAAGACCAAGGTCCCGCGCGCATACGACG 661  
Db 1804 CTCACCTCATGCTCCCAACCGCAGCGGTAAAGACCAAGGTCCCGCGCGCATACGACG 1863  
QY 662 CTCAGGGCTTACAAGGTGCTGGTGTCAACCCCTCCGCTGCTGTGCAACAATGGGCTTTGGTG 721  
Db 1864 CTCAGGGCTTACAAGGTGCTGGTGTCAACCCCTCCGCTGCTGTGCAACAATGGGCTTTGGTG 1923  
QY 722 CTTACATGTCAGAGCCCATAGGATTCCTAAACATCAGAGCTGGGTGAGGACAAATTA 781  
Db 1924 CTTACATGTCAGAGCCCATAGGATTCCTAAACATCAGAGCTGGGTGAGGACAAATTA 1983  
QY 782 CTACTGGCAGCGCCATCAGTATTCACCTACGGCAAGTTCCTTCCGACGCGCGGTGT 841  
Db 1984 CCACTGGCAGCGCCATCAGTATTCACCTACGGCAAGTTCCTTCCGACGCGCGGTGT 2043  
QY 842 CAGGGGTGCTTATGACATAATAATTTGTGAGAGTGCACCTCCACGATGCAACATCA 901  
Db 842 CAGGGGTGCTTATGACATAATAATTTGTGAGAGTGCACCTCCACGATGCAACATCA 901

Db 2044 CGGGGGCGCTTATGACATAATATTTGTGAGGACTGCCACTCCACGGATGCCACATCCA 2103  
Qy 902 TCTTGGGCAATGGCACTGCTCTTGACAAAGCAGAGACCGCGGGGCGAGACTGACTGTGC 961  
Db 2104 TCTTGGGCATCGGCACCTGCTCTTGACAAAGCAGAGACTGCGGGGCGAGACTGCTGTGTC 2163  
Qy 962 TCGCACCGCTACCCCTCCGGGCTCCGTCACGTGTGCCCCATCTTAACATCAGAGAGTTG 1021  
Db 2164 TCGCACCGCCACCCCTCCGGGCTCCGTCACGTGTGCCCCATCTTAACATCAGAGAGTTG 2223  
Qy 1022 CTCTGTCTACCTACCGGAGAGATCCCTTTTATGCAAGGCTATTCCCTTGAAGCAATTA 1081  
Db 2224 CTCTGTCCACACCGGAGAGATCCCTTTTATGCAAGGCTATTCCCTTGAAGCAATTA 2283  
Qy 1082 AGGGGGGAGACATCTCATCTTCTGCCACTCAAGAGAAAGTGCAGAGCTCCCGCAA 1141  
Db 2284 AGGGGGGAGACATCTCATCTTCTGCCACTCAAGAGAAAGTGCAGAGCTCCCGCAA 2343  
Qy 1142 AACTGGTCCGTTGGGCGTCAATGCGGTGCTTACTACCGGGCTTGATGTCCGTCA 1201  
Db 2344 AGCTGGTCCGATGGGCATCAATGCGGTGCTTACTACCGGGCTTGATGTCCGTCA 2403  
Qy 1202 TCCGACACAGTGGTGGCTGTGTCGTGGCAACTGACGCCCTCATGACCGGCTTTACCG 1261  
Db 2404 TCCGACACAGCGGAGTGTGTCGTGGCAACCGATGCCCTCATGACCGGCTTTACCG 2463  
Qy 1262 GCGACTTCGATTCGGTGTAGACTGCAACACGTGTGTCAACGACAGCTGACCTTCAGCC 1321  
Db 2464 GCGACTTCGATTCGGTGTAGACTGCAATACGTGTGTCAACGACAGCTGATTTAGCC 2523  
Qy 1322 TTGACCTTACCTTACCAATTGAGCAATCAGCTTCCCGAGGATGCTGTCTCCCGTACTC 1381  
Db 2524 TTGACCTTACCTTACCAATTGAGCAATCAGCTTCCCGAGGATGCTGTCTCCCGACTC 2583  
Qy 1382 AACGTCGGGTGAGACTGGCAGAGGAGGAGCAGCATCTACAGATTTGTGCAACCGGGG 1441  
Db 2584 AACGTCGGGCGAGACTGGCAGGAGGAGGAGCAGCATCTACAGATTTGTGCAACCGGGG 2643  
Qy 1442 AGCGTCTCTTGGCATGTTTGGACTCGTGTGCTCTGCGAGTGTATGAGCGGGTGTG 1501  
Db 2644 AGCGCCCTCCGGCATGTTTGGACTCGTGTGCTCTGCGAGTGTATGAGCGAGCTGTG 2703  
Qy 1502 CTGAGTATGACTTACGCCCGCGAGACACAGTTAGGCTACGAGCATACATGAACACCC 1561  
Db 2704 CTGAGTATGACTTACGCCCGCGAGACACAGTTAGGCTACGAGCGTACATGAACACCC 2763  
Qy 1562 CGGCACTTCCGTCGTCAGACCATCTTGAATTTGGGAGGCGCTTTTACGGGTCTCA 1621  
Db 2764 CGGCGCTTCCGTCGTCAGACCATCTTGAATTTGGGAGGCGCTTTTACAGGCGCTCA 2823  
Qy 1622 CCCATAGACGCCCACTTCTTATCCAGACAAAGCAGAGTGGGGAACCTTCCCTATC 1681  
Db 2824 CTATATAGATGCCCACTTCTTATCCAGACAAAGCAGAGTGGGAGAACCTTCTTACC 2883  
Qy 1682 TGTAGGCTACCAAGCACCGTGTGCGTAGAGCTCAAGGCCCTTCCCGGTCGTTGGAC 1741  
Db 2884 TGTAGGCTACCAAGCACCGTGTGCGTAGAGCTCAAGGCCCTTCCCGGTCGTTGGAC 2943  
Qy 1742 AGATGTGGAAGTGTGATCGCTCAAGCCACCTTCCATGGGCGCACCTTGTCTAT 1801  
Db 2944 AGATGTGGAAGTGTGATCGCTCAAGCCACCTTCCATGGGCGCACCTTGTCTAT 3003  
Qy 1802 ATAGACTGGGCGCTGTCCAGAAATGAAGTCAACCTTGGACCCAGTCAACCAAGTATATCA 1861  
Db 3004 ACAGACTGGGCGCTGTTCAGATGAATCACTTGAACGACCCAGTCAACCAATATATCA 3063  
Qy 1862 TGACATGTATGTCCGCTGAECTGGAGTGTGTCACAGTACTCTGGGTGCTCTGTTGGCGCG 1921  
Db 3064 TGACATGTATGTCCGCGACCTGGAGTGTGTCACGAGCACCTGGGTGCTCTGTTGGCGCG 3123  
Qy 1922 TTCTGGCTGTCTTGGCGCGTATTTGCCTATCCACAGGCTGGGTGCTATAGTAGGTAGGA 1981  
Db 3124 TCCTGGCTGTCTTGGCGCGTATTTGCCTGTCAACAGGCTGGGTGCTATAGTAGGTAGGG 3183

Qy 1982 TTGTCTTTGTCCGAAAGCCGGCAATCATACCCGACAGGAAGTCTCTTACCGGAGTTGC 2041  
Db 3184 TCGTCTTTGTCCGGAAGCCGGCAATCATACCTGACAGGAAGTCTCTTACCGAGATTGC 3243  
Qy 2042 ATGAAATGGAAGAGTGCT 2059  
Db 3244 ATGAGATGGAAGAGTGCT 3261

## RESULT 10

US-08-444-818-88  
; Sequence 88, Application US/08444818  
; Patent No. 6150087  
; GENERAL INFORMATION:  
; APPLICANT: Chien, David Y.  
; APPLICANT: Rutter, William J.  
; TITLE OF INVENTION: NANBV Diagnostics and Vaccines  
; NUMBER OF SEQUENCES: 777  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Chiron Corporation  
; STREET: 4560 Horton Street  
; CITY: Emeryville  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 94608-2916  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/444,818  
; FILING DATE:  
; CLASSIFICATION: 424  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/08/403,590  
; FILING DATE: 14-MAR-1995  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Harbin, Alisa A.  
; REGISTRATION NUMBER: 33,895  
; REFERENCE/DOCKET NUMBER: 0110.002  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (508)359-3876  
; TELEFAX: (508)359-3885  
; INFORMATION FOR SEQ ID NO: 88:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 8316 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: cDNA  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: 1..8316  
; US-08-444-818-88

Query Match 86.6%; Score 1784.4; DB 3; Length 8316;  
Best Local Similarity 91.7%; Pred. No. 0;  
Matches 1887; Conservative 0; Mismatches 171; Indels 0; Gaps 0;

Qy 2 TGGCGCTATCAGCGCTATGCCAGCAGACAAAGGGGCTTTTGGGATGCATAATCACCA 61  
Db 2735 TGGCGCCCATCAGCGGTAGCCCGCAGCAGACAGGGGCTCTTAGGTGCATATCACCA 2794  
Qy 62 GCTTGACCGCGGACAAACCAAGTGGAGGTGAGTTGAGATCGTGTCAACTCTG 121  
Db 2795 GCCTAACTGGCGGACAAACCAAGTGGAGGTGAGTCCAGATTGTGTCAACTCTG 2854  
Qy 122 CCAGACTTCTTGGCAACCTGATTAACCGGGTGTGTGGACTGTCTACATGGAGCCG 181  
Db 2855 CCACAACTTCTTGGCAACCTGATTAACCGGGTGTGTGGACTGTCTACATGGAGCCG 2914

182 GAACAAGACCAATTCCTGCTCACTAAGGGTCTCTGTTATCCAGATGTATACCAATGTGGACC 241  
Db  
2915 GAACGAGGACCAATCGCGTCAACCAAGGGTCTCTGTTATCCAGATGTATACCAATGTAGACC 2974  
Qy  
242 AAGACCTCGTAGGCTGCGCTCCCGCAAGGTGCGCTCATTAACACCAATGCACTTGGC 301  
Db  
2975 AAGACCTTGTGGGTGGCGCTCCGCAAGGTAGCGCTCATTTGACACACCTGCACTTGGC 3034  
Qy  
302 GCTCTCGGACCTTACCTGCTCAGAGGACCGCGATGTCATTCCTGTGCGCGGCGG 361  
Db  
3035 GCTCTCGGACCTTACCTGCTCAGAGGACCGCGATGTCATTCCTGTGCGCGGCGG 3094  
Qy  
362 GTGATGCGAGGGGAGCTCTCTTTGCGCCCGGCTATCTCTTACTTGAAGGCTCTCTCGG 421  
Db  
3095 GTGATGCGAGGGGAGCTCTCTGCTGCGCCCGGCTCTCTTCTACTTGAAGGCTCTCTCGG 3154  
Qy  
422 GAGGCCCTCTGCTGTCGCGCAGGACATCCGTAGGCAATTAAGAGCCGCGGTATGCA 481  
Db  
3155 GGGGTCCGCTGTTGTGCCCCCGGGGACCGCCGTGGGCAATTTAGGGCCGCGGTGTGCA 3214  
Qy  
482 CCGTGGAGTGGCTAAGGCGGTGGATCTCATCCCGGTAGAGCTTAGAGACAACCATGA 541  
Db  
3215 CCGTGGAGTGGCTAAGGCGGTGGATCTTATCCCTGTGGAGAACCTTAGAGACAACCATGA 3274  
Qy  
542 GGTCCCCGGTGTCTCAGACAACTCTCTCCCAAGCAGAGTGCCTCCAGAGCTACCAAGTGG 601  
Db  
3275 GGTCCCCGGTGTCTCAGAGTAACTCTCTCTCCCAAGTGTGCTCCAGAGCTTCCAGGTGG 3334  
Qy  
602 CCACCTGCAATGCTCCCAAGGCGGTAAAGAGCAACAAAGTCTCCGCGCGCATACGCGAG 661  
Db  
3335 CTCACTTCAATGCTCCCAAGGCGGTAAAGAGCAACAAAGTCTCCGCGGTGCAATATGCGAG 3394  
Qy  
662 CTCAAGGCTCAAGGTGCTGCTCTCAACCCCTCGTGTGCTGCAACAAAGGCGCTTGGTG 721  
Db  
3395 CTCAAGGCTCAAGGTGCTGCTCTCAACCCCTCGTGTGCTGCAACAAAGGCGCTTGGTG 3454  
Qy  
722 CTTACATGTCCAAGGCTCAAGGTGCTCTCAACCATCAGGACTGGGGGTGAGCAATTA 781  
Db  
3455 CTTACATGTCCAAGGCTCAAGGTGCTCTCAACCATCAGGACTGGGGGTGAGCAATTA 3514  
Qy  
782 CTACTGCGAGCCGATCAGCTATTCACCTACGGCAAGTTCCTTGGCGAGCGCGGTGTT 841  
Db  
3515 CCCTGCGAGCCGATCAGCTATTCACCTACGGCAAGTTCCTTGGCGAGCGCGGTGCT 3574  
Qy  
842 CAGGGGTGCTTATGACATAAATTTGTACAGAGTGCCACTCCAGGATGCAACATCCA 901  
Db  
3575 CGGGGGCGGTATGACATAAATTTGTACAGAGTGCCACTCCAGGATGCCACATCCA 3634  
Qy  
902 TCTTGGGCATTGGCACTGTCTTGACCAAGCAGAGACCGGGGGGAGAGCTGACTGTGC 961  
Db  
3635 TCTTGGGCATTGGCACTGTCTTGACCAAGCAGAGACTGGGGGGGAGAGCTGTTGTGC 3694  
Qy  
962 TCGCACCCTGCTACCTCGGGTCTCGTCACTGTGCCCCATCTTAACATGAGAGGTTG 1021  
Db  
3695 TCGCACCCTGCTACCTCGGGTCTCGTCACTGTGCCCCATCTTAACATGAGAGGTTG 3754  
Qy  
1022 CTCTGTCCACTACCGAGAGATCCCTTTTATGCAAGGCTATTCCTCTTGAAGCAATTA 1081  
Db  
3755 CTCTGTCCACTACCGAGAGATCCCTTTTATGCAAGGCTATTCCTCTTGAAGCAATTA 3814  
Qy  
1082 AGGGGGGAGACATCTCATCTTCTGCGCATCTCAAGAGAGTGGCGAGAGCTCGCGCAA 1141  
Db  
3815 AGGGGGGAGACATCTCATCTTCTGCTCATTAAGAGAGTGGCGAGAGCTCGCGCAA 3874  
Qy  
1142 AACTGTCGGTGGGGGTCAATGCGGTGCTTACTACCGCGCTTGTGATGTCCGTC 1201  
Db  
3875 AGCTGTGCAATTTGGGATCAATGCGGTGCTTACTACCGCGCTTGTGATGTCCGTC 3934  
Qy  
1202 TCCCGACCACTGCTGCTGCTGCGCAACTGACGCCCTCATGACCGGCTTACCG 1261  
Db  
3935 TCCCGACCACTGCTGCTGCTGCGCAACTGACGCCCTCATGACCGGCTTACCG 3994  
Qy  
1262 GCGACTTCGATTCGGTGATAGACTGCAACACGTGTGTACCCAGACAGTCTCAGCC 1321

Db  
3995 GCGACTTCGACTCGGTGATAGACTGCAATACGTGTGTACCCAGACAGTCTGATTCAGCC 4054  
Qy  
1322 TTGACCCCTACTCTTACCATTTGAGACAAATCAGCTTCCCGAGATGCTGTCTCCGTACTC 1381  
Db  
4055 TTGACCCCTACTCTTACCATTTGAGACAAATCAGCTTCCCGAGATGCTGTCTCCCGACTC 4114  
Qy  
1382 AACCTCGGGGTAGGACTGCGCAGAGGGAAGCCAGGCAATCTACAGATTTGTGGACCGGGG 1441  
Db  
4115 AACCTCGGGGTAGGACTGCGCAGAGGGAAGCCAGGCAATCTACAGATTTGTGGACCGGGG 4174  
Qy  
1442 AGCGTCTCTTCTGGCAATTTTGACTCGTCTCTCTGCGAGTCTATGACGCGGTTGTG 1501  
Db  
4175 AGCGTCTCTTCTGGCAATTTTGACTCGTCTCTCTGCGAGTCTATGACGCGGTTGTG 4234  
Qy  
1502 CTTGTGATGAGCTTAGCCCGCGGACCAACAGTGTAGGCTACGAGCATATGAAACACC 1561  
Db  
4235 CTTGTGATGAGCTTAGCCCGCGGACCAACAGTGTAGGCTACGAGCATATGAAACACC 4294  
Qy  
1562 CCGGACTTCCCGTGTGCCAAGACCAATCTTTGAAATTTTGGAGGCGTCTTTACGGGTCTCA 1621  
Db  
4295 CCGGGCTTCCCGTGTGCCAAGACCAATCTTTGAAATTTTGGAGGCGTCTTTACGGCTCTCA 4354  
Qy  
1622 CCCACATAGACGCCCACTTCTTATCCAGACAAAGCAGAGTGGGGAACCTTCCCTATC 1681  
Db  
4355 CTCATATAGATGCCCACTTCTTATCCAGACAAAGCAGAGTGGGGAACCTTCCCTATC 4414  
Qy  
1682 TGGTAGCTTACCAAGCCACCGTGTGGCTAGAGCTCAAGCCCTCCCGTGTGGGACC 1741  
Db  
4415 TGGTAGCTTACCAAGCCACCGTGTGGCTAGAGCTCAAGCCCTCCCGTGTGGGACC 4474  
Qy  
1742 AGATGTGGAAGTGTGTTGATCCGCTCAAGCCACCGTGTGGCTCAAGCCCTCCCGTGTGGGACC 1801  
Db  
4475 AGATGTGGAAGTGTGTTGATCCGCTCAAGCCACCGTGTGGCTCAAGCCCTCCCGTGTGGGACC 4534  
Qy  
1802 ATAGACTGGGCTGTCTCAGAAATGAAGTCACTGACGACCCAGTCAACCAAGTATATCA 1861  
Db  
4535 ACAGACTGGGCTGTCTCAGAAATGAAGTCACTGACGACCCAGTCAACCAAGTATATCA 4594  
Qy  
1862 TGACATGATGTGGCTGACCTGGAGTGTGTCAGAGTACCTGGGCTGGTGGCGGG 1921  
Db  
4595 TGACATGATGTGGCTGACCTGGAGTGTGTCAGAGTACCTGGGCTGGTGGCGGG 4654  
Qy  
1922 TTCTGGCTCTTGGCGGCTATTCCTATCCAGGCTGCGTGTGCTAGTAGGTAGGA 1981  
Db  
4655 TCTTGGCTCTTGGCGGCTATTCCTATCCAGGCTGCGTGTGCTAGTAGGTAGGA 4714  
Qy  
1982 TTGTCTTGTCCGGAAGCCGCAATCATACCGACAGGGAAGTCTCTTACCGGGAGTTG 2041  
Db  
4715 TCGTCTTGTCCGGAAGCCGCAATCATACCTGACAGGGAAGTCTCTTACCGAGAGTTG 4774  
Qy  
2042 ATGAATGGAAGTGTCT 2059  
Db  
4775 ATGAGATGGAAGTGTCT 4792

RESULT 11  
US-08-444-818-137  
; Sequence 137, Application US/08444818  
; Patent No. 6150087  
; GENERAL INFORMATION:  
; APPLICANT: Chien, David Y.  
; APPLICANT: Rutter, William J.  
; TITLE OF INVENTION: NANEV Diagnostics and Vaccines  
; NUMBER OF SEQUENCES: 777  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Chiron Corporation  
; STREET: 4560 Horton Street  
; CITY: Emeryville  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 94608-2916  
; COMPUTER READABLE FORM:



MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/444,818  
FILING DATE:  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/403,590  
FILING DATE: 14-MAR-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Harbin, Alisa A.  
REGISTRATION NUMBER: 33,895  
REFERENCE/DOCKET NUMBER: 0110.002  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (508)359-3876  
TELEFAX: (508)359-3885  
INFORMATION FOR SEQ ID NO: 137:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 8987 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 1..8985

US-08-444-818-137

Query Match 86.6%; Score 1784.4; DB 3; Length 8987;

Best Local Similarity 91.7%; Pred. No. 0;

Matches 1887; Conservative 0; Mismatches 171; Indels 0; Gaps 0;

Qy	2	TGGCGCCTATCAGCGCCTATCCAGCAGACAAAGGGCCCTTTGGGTGATATACCA	61
Db	3077	TGGCGCCCATCACGGCGTACCCAGCAGACAAAGGGCCCTCTAGGGTGCATATACCA	3136
Qy	62	GCTTGACCGGCGGACAAACAGGTGGAGGTGAGGTTTCAGATCGTGCACTGCTG	121
Db	3137	GCCTAATCGGCGGAGCAAAACAGTGGAGGTGAGTCCAGATTGTGCACTGCTG	3196
Qy	122	CCCAGACTTCTTGCGCAACTGCAATTAACGGGGTGTGTGGACTGTCTACCATGAGCCG	181
Db	3197	CCCAAACTTCTTGGCAACGTGCATCAATGGGTGTGTGACTGTCTACCAAGGGCCG	3256
Qy	182	GAACAGGACCATTGGCTACCTAAGGTCCTGTATCCAGATGTACCAATGTGAC	241
Db	3257	GAACGAGACCATCGCGTCACCCAAAGGTCTGTGACTGTCTATCCAGATGTATACCA	3316
Qy	242	AAGACCTGTAAGCTGGCCGCTCCCAAGGTGCCGCTCATTAACACCATGCACTTGG	301
Db	3317	AAGACCTTGTGGCTTGGCCGCTCCGAGGTAGCCGCTCATTAACACCATGCACTTGG	3376
Qy	302	GCTCCTCGGACCTTTACCTGTGTACAGGACGCGCGATGTCTTCTGTGCGCGACGGG	361
Db	3377	GCTCCTCGGACCTTTACCTGTGTACAGGACGCGCGATGTCTTCTGTGCGCGCGGG	3436
Qy	362	GTGATGGAGGGGAGCGCTGTCTTGGCCCGGCGCTATCTTCTTGAAGGCTCTCTGG	421
Db	3437	GTGATAGCAGGGGAGCGCTGTCTTGGCCCGGCGCTATCTTCTTGAAGGCTCTCTGG	3496
Qy	422	GAGCGCTCTGCTGGCCCGCAGACATGCGGTAGGCATATTACAGACCGCGGTATGCA	481
Db	3497	GGGGTCCGCTGTGTGGCCCGGCGGACGCGGTGGGCATATTAGGCGCGCGGTATGCA	3556
Qy	482	CCCGTGGAGTGGCTAAGCGGTGGACTTTCATCCCGGTAGAGGCTTAGAGACCAACATGA	541
Db	3557	CCCGTGGAGTGGCTAAGCGGTGGACTTTCATCCCGGTAGAGGCTTAGAGACCAACATGA	3616
Qy	542	GGTCCCGGTGTTCTAGACAACTCTCCCGCAGCAGATGCCCGCAGAGTACCAAGTGG	601
Db	3617	GGTCCCGGTGTTCTAGCAGTAACCTCTCTCCAGCAGTAGTCCCGCAGAGCTTCCAGGTGG	3676

Qy	602	CCACCTTGCTGCTCCACCGGCGAGCGGTAAAGACACCAAGGTCCCGCGCATACGCGAG	661
Db	3677	CTCACCTCCATGCTCCACAGGCGAGCGGCAAAACCAAGGTCCCGGCTCATATGCGAG	3736
Qy	662	CTCAGGGCTACAGGTGCTGCTCAACCCCTCGGTTGTCGCAACAATGGGCTTTGGTG	721
Db	3737	CTCAGGGCTATAAGGTGCTAGTACTCAACCCCTCTGTTGTCGAACACTGGGCTTTGGTG	3796
Qy	722	CTTACATGTCCAAGGCCATGGGATTGATCTTAACATCAGGACTGGGTGAGGACAATTA	781
Db	3797	CTTACATGTCCAGGCTCATGGGATCGATCTTAACATCAGGACCGGGGTGAGACAATTA	3856
Qy	782	CTACTGCGAGCCCGATCACGTAATTCACCTACGCAAGTTCTCTTGGCGACCGCGGTGT	841
Db	3857	CCACTGGCAGCCCATCACTACTCCACTAGCGCAAGTTCTCTTGGCGACCGCGGTGT	3916
Qy	842	CAGGGGTGCTTATGACATTAATTTGTACGAGTGCCTCCACCGATCAACATCA	901
Db	3917	CGGGGGCGCTTATGACATTAATTTGTACGAGTGCCTCCACCGATCAACATCA	3976
Qy	902	TCCTGGGCATTTGGCAGTGTCTTGACCAAGCAGAGACCGGGGGCGAGACTGACTGTGC	961
Db	3977	TCCTGGGCATCGGCATGTCTTGACCAAGCAGAGACTGCGGGGGCGAGACTGTTGTGC	4036
Qy	962	TCGCCACCGCTTACCCCTCCGGGCTCCGTCACTGTGCCCATCTCTTAACATCAGGAGGTG	1021
Db	4037	TCGCCACCGGCACCCCTCCGGGCTCCGTCACTGTGCCCATCTCCCAACATCAGGAGGTG	4096
Qy	1022	CTCTGTCCACTACCGGAGAGATCCCTTTTATGCAAGGCTATTTCCCTTGTGAAGCAATTA	1081
Db	4097	CTCTGTCCACTACCGGAGAGATCCCTTTTATGCAAGGCTATTTCCCTTGTGAAGTAATCA	4156
Qy	1082	AGGGGGGAGACATCTCATCTTCTGCCACTCAAGAGAGAGTGCAGAGCTGCCCGCAA	1141
Db	4157	AGGGGGGAGACATCTCATCTTCTGCCACTCAAGAGAGAGTGCAGAGCTGCCCGCAA	4216
Qy	1142	AACTGGTGGGCTCAATGCCGCTTACTACCGCGCTTGTGATGTGCCGTCA	1201
Db	4217	AGCTGGTGGGCTCAATGCCGCTTACTACCGCGCTTGTGATGTGCCGTCA	4276
Qy	1202	TCCGACACAGTGTGAGTGTGCTGTGGCAACTGACGCCCTCATGACCGGCTTTACCG	1261
Db	4277	TCCGACACAGCGGCTGTTGTGCTGTGGCAACCGATGCCCTCATGACCGGCTATACCG	4336
Qy	1262	GCGACTTGGATTCGGTATAGACTGCAACAGTGTGTCAACAGACAGTGCAGCTTCAGCC	1321
Db	4337	GCGACTTGGATTCGGTATAGACTGCAATACGTGTGTCAACAGACAGTGCAGCTTCAGCC	4396
Qy	1322	TTGACCTTACCTTCAACATTTAGACATCACTGCTTCCCGAGGATGCTGTCTCCGCTACTC	1381
Db	4397	TTGACCTTACCTTCAACATTTAGACATCACTGCTTCCCGAGGATGCTGTCTCCGCTACTC	4456
Qy	1382	AAGTCCGGGTAGGACTGGCAGAGGAGGACGAGCATCTACAGATTTGTGGCAGCGGGG	1441
Db	4457	AAGTCCGGGTAGGACTGGCAGAGGAGGAGGACGAGCATCAACAGATTTGTGGCAGCGGGG	4516
Qy	1442	AGCGTCTTCTGGCATGTTTGAATCGTCTCTCTGCGAGTGTATGACCGGGGTGTG	1501
Db	4517	AGCGTCTTCTGGCATGTTTGAATCGTCTCTCTGCGAGTGTATGACCGGGGTGTG	4576
Qy	1502	CTTGGTATGAGCTTACCGCCCGGAGACCAAGTTAGGCTACGAGCATACATGAACACCC	1561
Db	4577	CTTGGTATGAGCTACCGCCCGGAGACTACAGTTAGGCTACGAGCGTACATGAACACCC	4636
Qy	1562	CGGACTTCCCGTGTGCCAAGACCATCTTGAATTTTGGGAGGCGCTTTTACGGGTCTCA	1621
Db	4637	CGGACTTCCCGTGTGCCAAGACCATCTTGAATTTTGGGAGGCGCTTTTACAGGCTCTCA	4696
Qy	1622	CCACATAGAGCGCCACTTCTCTATCCAGACAAAGCAGAGTGGGGAACCTTCCCTATC	1681
Db	4697	CTCATATAGATGCCCACTTCTATCCAGACAAAGCAGAGTGGGGAACCTTCTCTTACC	4756

Qy	1682	TGTTAGCGTACCAAGCACCGTGTGCCTAGAGCTCAAGCCCTTCCCCCTGTGTTGGAGCC	1741
Db	4757	TGTTAGCGTACCAAGCACCGTGTGCCTAGGGCTCAAGCCCTTCCCCCATCTGTGGAGCC	4816
Qy	1742	AGATGTGGAGTGTTCATCCGTCTCAAGCCCCCTTCCATGGCGCAACACCTCTGCTAT	1801
Db	4817	AGATGTGGAGTGTTCATTCGCTCAAGCCCCCTTCCATGGCGCAACACCTCTGCTAT	4876
Qy	1802	ATAGACTGGCGCTGTCCAGAATGAAAGTCACTTCAGCGCACCCAGTCACCAAGTATATCA	1861
Db	4877	ACAGACTGGCGCTGTTCAGATGAAATCACCTTCAGCGCACCCAGTCACCAATATATCA	4936
Qy	1862	TGACATGTATGTCCGGCTGAACCTGAGAGTGTGTACAGATACTGGGTGCTCGTTGGCGGCG	1921
Db	4937	TGACATGCATGTCCGGCGACCTTGAGGTGTGTACAGACACTGGGTGCTCGTTGGCGGCG	4996
Qy	1922	TTCTGGGCTGCTTTGGCGCGTATTTGCTATTCACACAGGCTGGGTGTATAGTAGGTAGGA	1981
Db	4997	TCCTTGGCTGCTTTGGCGCGTATTTGCTGTCAACAGGCTCGGTGGTCAATAGTTGGCGAGGG	5056
Qy	1982	TTGTCTTTGTCCGGAAGACCGGCAATCATACCCGACAGGGAAGTCTCTTACCGGAGTTTCG	2041
Db	5057	TCGTCTTTGTCCGGAAGACCGGCAATCATACCTGCACAGGGAAGTCTCTTACCGAGAGTTTCG	5116
Qy	2042	ATGAAATGGAAGAGTGTCT	2059
Db	5117	ATGAGATGGAAGAGTGTCT	5134

```

RESULT 12
US-08-444-818-122
; Sequence 122, Application US/08444818
; Patent No. 6150087
; GENERAL INFORMATION:
; APPLICANT: Chien, David Y.
; APPLICANT: Rutter, William J.
; TITLE OF INVENTION: NANV Diagnostics and Vaccines
; NUMBER OF SEQUENCES: 777
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Chiron Corporation
; STREET: 4560 Horton Street
; CITY: Emeryville
; STATE: CA
; COUNTRY: USA
; ZIP: 94608-2916
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentn Release #1.0, Version #1.30
; CURRENT APPLICATION NUMBER: US/08/444,818
; FILING DATE:
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/403,590
; FILING DATE: 14-MAR-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Harbin, Alisa A.
; REGISTRATION NUMBER: 33,895
; REFERENCE/DOCKET NUMBER: 0110.002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (508)359-3876
; TELEFAX: (508)359-3885
; INFORMATION FOR SEQ ID NO: 122:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 9185 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cdna
US-08-444-818-122

```

Query Match	86.6%	Score 1784.4	DB 3	Length 9185
Best Local Similarity	91.7%	Pred. No. 0		
Matches 1887	Conservative	0	Mismatches 171	Indels 0
Qy	2	TGGCGCTATCACGGCCTATGCCAGACAGACAGGGGCGCTTTTGGGATGCATTAATCACCA	61	
Db	3396	TGGGCGCCATCACGGCGTAGCCCCAGACAGACAGGGGCGCTCTTAGGTGCATAATCACCA	3455	
Qy	62	GCTTGACCGCGCGGAGCAAAAACACAGGTGAGGGTGAGGTTAGATCTGTCAACTGCTG	121	
Db	3456	GCCTAACTGCCCGGAGCAAAAACAAATGGAGGGTGAGGTTCAGATTTGTCTCAACTGCTG	3515	
Qy	122	CCAGACCTTTCTTGGCAACCTGCATTAAACGGGGTGTTGGACATGCTCTACCATGGAGCGG	181	
Db	3516	CCCAACCTTCTGGGCACACGTGATCAATGGGGTGCTGGACATGCTCTACACGGGGCGG	3575	
Qy	182	GAACAAGGACCATTTGGGTACACTTAAGGGTCCTGTTATCCAGATGTGTACACCATGTGGACC	241	
Db	3576	GAACGAGGACCATCGGCTACCCCAAGGGTCCTGTCTATCCAGATGTGTATACCAATGTAGACC	3635	
Qy	242	AAGACCTGTAGGTGGCGCGCTCCCCAAGGTGCCGCTCATTTAAACACCATGCATTGCG	301	
Db	3636	AAGACCTTGTGGGTGGCGCGCTCCGCAAGGTAGCGCTCATTTGACACCTGTCACATTGCG	3695	
Qy	302	GCTCCTCGGACCTTTTACCTGGTACAGGAGCACGCCGATGTCTATCTCTGTGCGCGGCGG	361	
Db	3696	GCTCCTCGGACCTTTTACCTGGTCAAGGAGCACGCCGATGTCTATCTCGTCGCGCGGCGG	3755	
Qy	362	GTGATGCGAGGGGACGCTGCTTTGGCGCGGCGCTATCTTACTTGTGAAGGCGCTCCTCGG	421	
Db	3756	GTGATAGACGGGCGCGCTGCTGCGCGCGGCGGCTATCTCTACTTTGAAAGGCTCCTCGG	3815	
Qy	422	GAGGCCCTCTGCTGTGCCCGCAGGACATGCCGTGAGGCATATTCAGAGCGCGCGGTATGCA	481	
Db	3816	GGGGTCCGCTGTGTGCCCGCGGGGACGCGGTGGGCATATTTAGGGCGCGGTGTGCA	3875	
Qy	482	CCCGTGAGTGCGCTAAGCGGGTGGACTTCAATCCCGGTAGAGGTTAGAGCAACCATGA	541	
Db	3876	CCCGTGAGTGCGGTTAAGCGGGTGGACTTTATCCCTGTGGAGAACCTTAGAGCAACCATGA	3935	
Qy	542	GGTCCCAGGTGTTCTCAGACAACTCTCTCCCAACAGCAGTGCCCGCAGACTACCAAGTGG	601	
Db	3936	GGTCCCAGGTGTTCCAGGATAACTCTCTTCCACAGTAGTGTGCCCGAGCTTCCAGGTGG	3995	
Qy	602	CCCACTGCATGCTCCACCGGAGGGTTAAGAGCACCAAGGTCCCGGCGCATACGCAG	661	
Db	3996	CTCACTCCATGCTCCACAGGAGCGGCAAAAGCACCAAGGTCCCGGCTGCATATGCAG	4055	
Qy	662	CTCAGGGCTACAGGTGCTGGTGCTCAACCCCTCCGTTGTGCAACAATGGGCTTTGGTG	721	
Db	4056	CTCAGGGCTATAAGGTGCTAGTACTCAACCCCTCTGTGTGTGCAACTGGGCTTTGGTG	4115	
Qy	722	CTTACATGTCCAAGGCCCATGGGATTGATCCTTAACATCAGGACTGGGGGTGAGACAATTA	781	
Db	4116	CTTACATGTCCAAGGCTCATGGGATCGATCCTTAACATCAGGACCGGGGTGAGAACATTA	4175	
Qy	782	CTACTGGGAGGCCGATACGATATTCACACTACGGCAAGTTCTTTCGCAAGGGGGTGT	841	
Db	4176	CCACTGGCAGCCCATACGCTACTCCACCTACGGCAAGTTCTTTCGCGAGCGGGGGTGCT	4235	
Qy	842	CAGGGGTGCTTTATGACAATAATTTGTGACAGTGCCTACTCCAGGATGCACCATCCA	901	
Db	4236	CGGGGGCGGTTATGACATAATAATTTGTGACAGTGCCTACTCCACGATGCCACATCCA	4295	
Qy	902	TCTTTGGGACATGGCACTGTCTTTGACCAAGCAGAGACGGGGGGCGAGACTGACTGTGC	961	
Db	4296	TCITTGGGCATCGGCACTGTCTTTGACCAAGCAGAGACTGGGGGGCGAGACTTGTGTGTC	4355	
Qy	962	TCGCCACCGTACCCCTCCGGGCTCCGCTCACTGTGCCCACTCTCTAACTCGAGAGGTG	1021	
Db	4356	TCGCCACCGCACCCCTCCGGGCTCCGCTCACTGTGCCCACTCCCACTCGAGAGGTTG	4415	
Qy	1022	CTCTGTCCACTACCGGAGATCCCCCTTTTATGGCAAGGCTATTTCCTCTGACGAATTA	1081	

```
Db 4416 CTCTGTCCACACCGAGAGATCCCTTTTACGGCAAGGCTATCCCCCTCGAAGTAATCA 4475
Qy 1082 AGGGGGGAGACATCTCATCTTCTGCCACTCAAAAGAAAGTGGGAGAGCTCGCCGCAA 1141
Db 4476 AGGGGGGAGACATCTCATCTTCTGCCACTCAAAAGAAAGTGGGAGAGCTCGCCGCAA 4535
Qy 1142 AACTGGTCGGCTTGGCGGTCAATCGGCGGTCTTACTACCGGGCTTGATGTGCGGTCA 1201
Db 4536 AGCTGGTCGCAATGGGCAATCAATCGGCGGTCTTACTACCGGGCTTGATGTGCGGTCA 4595
Qy 1202 TCCGACACAGTGTGACGCTGCTGCTGCGGCAACTGACGCCCTCATGACCGGCTTTACCG 1261
Db 4596 TCCGACACAGCGGAGATGTTGCTGCTGCGGCAACCGATGCCCTCATGACCGGCTATACCG 4655
Qy 1262 GCGACTTCGATTCCGTTGATAGACTGCAACACGCTGTGTCAACCAGACAGTCCAGCTTCAGCC 1321
Db 4656 GCGACTTCGACTCGGTGATAGACTGCAATACGCTGTGTCAACCAGACAGTCCAGCTTCAGCC 4715
Qy 1322 TTGACCTTACCTTACCAATTGAGCAATCAACGCTTCCCGAGGATGCTGTCTCCCGTACTC 1381
Db 4716 TTGACCTTACCTTACCAATTGAGCAATCAACGCTTCCCGAGGATGCTGTCTCCCGGACTC 4775
Qy 1382 AACGTCGGGTAGACTGGGAGAGGAGCCAGGATCTACAGATTTGTGGCACCGGGG 1441
Db 4776 AACGTCGGGGCAGGACTGGCAGGGGAGCCAGGATCTACAGATTTGTGGCACCGGGG 4835
Qy 1442 AGCGTCCTTCTGGCATGTTTGACTCGCTGCTGCTCTGCGAGTGTATGACCGCGGTGTG 1501
Db 4836 AGCGCTTCTCGGCATGTTTGACTCGCTGCTGCTCTGCGAGTGTATGACCGCGGTGTG 4895
Qy 1502 CTTGGTATGACTTACCGCCCGGAGACACAGTTAGGCTACGAGCATACATGAACACCC 1561
Db 4896 CTTGGTATGACTTACCGCCCGGAGACACAGTTAGGCTACGAGCGTACATGAACACCC 4955
Qy 1562 CGGACTTCCCGTGTGCCAAGACCATCTTGAATTTTGGAGGGGCTTTTACGGGTCTCA 1621
Db 4956 CGGGGCTTCCCGTGTGCCAAGACCATCTTGAATTTTGGAGGGGCTTTTACAGGCGCTCA 5015
Qy 1622 CCACATAGACGCCCATCTTCTATCCAGACAAAGCAGAGTGGGAAACCTTCCCTATC 1681
Db 5016 CTATATATAGTCCCATCTTCTATCCAGACAAAGCAGAGTGGGAGAACCTTCTTACC 5075
Qy 1682 TGGTAGCGTACCAAGCCACCGTGTGCGCTAGAGCTCAAGCCCTTCCCGGTCGTGGGACC 1741
Db 5076 TGGTAGCGTACCAAGCCACCGTGTGCGCTAGGCTCAAGCCCTTCCCGGTCGTGGGACC 5135
Qy 1742 AGATGTGAAGTGTGATCCGCTCAAGCCACCGCTCCATGGGCGCAACCTCTGCTAT 1801
Db 5136 AGATGTGAAGTGTGATCCGCTCAAGCCACCGCTCCATGGGCGCAACCTCTGCTAT 5195
Qy 1802 ATAGACTGGGCGCTGTCCAGAAATGAAGTCAAGTCAAGCCAGTCAAGTCAAGTATATCA 1861
Db 5196 ACAGACTGGGCGCTGTTCAGAAATGAATCAAGTCAAGCCAGTCAAGTCAAGTATATCA 5255
Qy 1862 TGACATGTATGTCGGCTGACCTGGAGGTGCTCAAGAGTACCTGGGCTGCTGCTGGCGCG 1921
Db 5256 TGACATGTATGTCGGCGACCTGGAGGTGCTCAAGAGTACCTGGGCTGCTGCTGGCGCG 5315
Qy 1922 TTCTGGCTGCTTGGCGCGGTATGCTTATCCAGAGTGGGTGCTGATAGTATAGTA 1981
Db 5316 TCCTGGCTGCTTGGCGCGGTATGCTTATCCAGAGTGGGTGCTGATAGTATAGTA 5375
Qy 1982 TTGTCTTGTCCGGAAGCGCAATCATACCCGACGAGGAGTCTCTTACCGGAGTTCG 2041
Db 5376 TCGTCTTGTCCGGAAGCGGCAATCATACCTGACAGGAAAGTCTCTTACCGAGATTTCG 5435
Qy 2042 ATGAATGGAAGAGTGTCT 2059
Db 5436 ATGAGATGGAAGAGTGTCT 5453
```

RESULT 13

```
US-08-444-818-123/c
; Sequence 123, Application US/08444818
; Patent No. 6150087
; GENERAL INFORMATION:
; APPLICANT: Chien, David Y.
; APPLICANT: Rutter, William J.
; TITLE OF INVENTION: NANBV Diagnostics and Vaccines
; NUMBER OF SEQUENCES: 777
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Chiron Corporation
; STREET: 4560 Horton Street
; CITY: Emeryville
; STATE: CA
; COUNTRY: USA
; ZIP: 94608-2916
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/444,818
; FILING DATE:
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/403,590
; FILING DATE: 14-MAR-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Harbin, Alisa A.
; REGISTRATION NUMBER: 33,895
; REFERENCE/DOCKET NUMBER: 0110.002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (508)359-3876
; TELEFAX: (508)359-3885
; INFORMATION FOR SEQ ID NO: 123:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 9185 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; ANTI-SENSE: YES
; US-08-444-818-123
```

```
Query Match 86.6%; Score 1784.4; DB 3; Length 9185;
Best Local Similarity 91.7%; Pred. No. 0;
Matches 1887; Conservative 0; Mismatches 171; Indels 0; Gaps 0;

Qy 2 TGGCGCTTATCAGCGCTATGCCAGCAGACAAGGGGCTTTTGGGATGCAATAATCACCA 61
Db 5790 TGGCGCTTATCAGCGCTATGCCAGCAGACAAGGGGCTTTTGGGATGCAATAATCACCA 5731
Qy 62 GCTTGACCGGCGGGAACAAAACAGGTGGAGGTTCAGATCGTGTCAACTGCTG 121
Db 5730 GCCTTAACGTGCGCGGACAAAACCAAGTGGAGGTGAGGTCCAGATTGTGTCAACTGCTG 5671
Qy 122 CCACAGACTTCTTGGCAACCTGCAATTAACGGGGTGTGTGGACTGTGTACCATGGAGCG 181
Db 5670 CCACAACTTCTTGGCAACCTGCAATTAACGGGGTGTGTGGACTGTGTACCAACGGGCG 5611
Qy 182 GAACAAGGACCAATGCGTCACTTAAGGGTCTGTATTCCAGATGATACCAATGTGAGCC 241
Db 5610 GAACAGGACCAATGCGTCACTTAAGGGTCTGTGTATCCAGATGATATACCAATGTAGACC 5551
Qy 242 AAGACCTCGTAGGTGCGCGCTCCCAAGGTGCGGCTCATTAACACCATGCACTTGGC 301
Db 5550 AAGACCTGTGGGTGCGCGCTCCCAAGGTGAGCTCATTAACACCATGCACTTGGC 5491
Qy 302 GCTCCTCGGACCTTACCTGTGTACAGGACGACCGGATGATCTCTGTGTGCGCGGAGCGG 361
Db 5490 GCTCCTCGGACCTTACCTGTGTACAGGACGACCGGATGATCTCTGTGTGCGCGGAGCGG 5431
Qy 362 GTGATGGAGGGGAGCGCTGCTTTTCGCCCCCGGCTATCTTCTTAAAGGCTCTCTCGG 421
```

Db 5430 GTGATAGAGGGGAGCCTGCTGTGCGCCCGGCCATTTCTACTTTGAAAGGCTCTCTCGG 5371  
Qy 422 GAGGCGCTCTGCTGTGCGCCCGCAGGACATGCGGTAGGCATTTTCAGAGCCGCGGTATGCA 481  
Db 5370 GGGGTCCGCTGTGTGCGCCCGCGGGGACGCCCTGGGCAATTTTAGGGCCGCGGTGTGCA 5311  
Qy 482 CCCGTGAGGTGGCTTAAGCGCGTGGACTTTCATCCCGGTAGAGAGCTTAGAGACAACCATGA 541  
Db 5310 CCCGTGAGGTGGCTTAAGCGCGTGGACTTTCATCCCTGTGGAGAACCTTAGAGACAACCATGA 5251  
Qy 542 GGTCCCGGTGTTCTCAGACAACCTCTCCACAGCAGTGCCTCCAGAGCTTACCAAGTGG 601  
Db 5250 GGTCCCGGTGTTACAGGATAAATCTCTCTCCACAGTAGTGCCTCCAGAGCTTCCAGGTGG 5191  
Qy 602 CCACCTGATGCTCCACCGGAGCGGTAAGAGCACAAGTTCCTCCGCGCGGATAGCGAG 661  
Db 5190 CTCACTTCCATGTCTCCACAGGAGCGGCAAAAGCACCAAGGTCCCGGCTGCATATGCAG 5131  
Qy 662 CTAGGGCTACAAGGTGCTGCTCAACCCCTCCGTTGCTGCAACAATGGGCTTTGGTG 721  
Db 5130 CTAGGGCTAAGGTGCTAGTACTCAACCCCTCTGTTGCTGCAACAATGGGCTTTGGTG 5071  
Qy 722 CTTACATGTCAAAGGCGCATGGGATTTGATCTTAACATCAGGACTGGGGTGAGGACAATTA 781  
Db 5070 CTTACATGTCAAAGGCTCATGGGATCGATCTTAACATCAGGACCGGGGTGAGAACAAATTA 5011  
Qy 782 CTACTGGCAGCCGATCAGCTATTCACCTACGGCAAGTTCCTTGGCAGCGGGGTGTT 841  
Db 5010 CCACCTGACGCCCATCACGTACTTCCACCTACGGCAAGTTCCTTGGCAGCGGGGTGCT 4951  
Qy 842 CAGGGGTGCTATGACATAAATTTGTGACGAGTGCCACTCCACGGATGCAACATCCA 901  
Db 4950 CGGGGGCGGTTAAGACATAAATTTGTGACGAGTGCCACTCCACGGATGCCACATCCA 4891  
Qy 902 TCTTGGSCATTGGCACTGTCTTTGACCAAGCAGAGACCGCGGGGGGAGACTGACTGTGC 961  
Db 4890 TCTTGGSCATCGCACTGTCTTTGACCAAGCAGAGACTGCGGGGGCGAGACTGTTGTGC 4831  
Qy 962 TCGCACCGGTACCGCTCCGGGCTCCGTCACTGTGCCCCATCTTAACATCAGGAGGTG 1021  
Db 4830 TCGCACCGGACCGCTCCGGGCTCCGTCACTGTGCCCCATCTTAACATCAGGAGGTG 4771  
Qy 1022 CTCTGTCCACTACCGGAGAGATCCCTTTTATGCAAGGCTATTCCTTTGAAAGCAATTA 1081  
Db 4770 CTCTGTCCACCGGAGAGATCCCTTTTATGCAAGGCTATTCCTTTGAAAGCAATTA 4711  
Qy 1082 AGGGGGGAGACATCTCATCTTCTGCCACTCAAAGAGAAAGTGGCAGAGCTCGCGCAA 1141  
Db 4710 AGGGGGGAGACATCTCATCTTCTGCCACTCAAAGAGAAAGTGGCAGAGCTCGCGCAA 4651  
Qy 1142 AACTGTGCGGTGGGCGTCAATGCGGTGCTTACTACCGCGGCTTGTATGTGCGTCA 1201  
Db 4650 AGTGTGCGATTTGGGCAATCAATGCGGTGCTTACTACCGGCTTGTACGTTGCGTCA 4591  
Qy 1202 TCCGACCACTGTGTGCTGTGTCGCACTGACGCCCTCATGACCGGCTTACCG 1261  
Db 4590 TCCGACCACTGTGTGCTGTGTCGCACTGACGCCCTCATGACCGGCTTACCG 4531  
Qy 1262 GCACATTCGATTCGGTGATAGACTGCAACACGTTGTTCACCCAGACAGTGCATTCAGCC 1321  
Db 4530 GCACATTCGATTCGGTGATAGACTGCAACACGTTGTTCACCCAGACAGTGCATTCAGCC 4471  
Qy 1322 TTGACCCCTACCTTCACTTTGAGACAAATCAAGCTTCCCGAGGATGTGTCTCCCGTACTC 1381  
Db 4470 TTGACCCCTACCTTCACTTTGAGACAAATCAAGCTTCCCGAGGATGTGTCTCCCGACTC 4411  
Qy 1382 AACGTGGGGTAGGACTGGCAGAGGGAAGCAGGCACTTACACATTTGTGGCACCGGGG 1441  
Db 4410 AACGTGGGGTAGGACTGGCAGAGGGAAGCAGGCACTTACACATTTGTGGCACCGGGG 4351  
Qy 1442 AGCGTCTCTTGGCATGTTTGAATCTGCTCTGCTGCGAGTGTCTATGACGCGGGTGTG 1501

Db 4350 AGCGCCCTCCGCGCATGTTTCGACTGCTCGTCTCTGTGAGTGTCTATGACGAGGCTGTG 4291  
Qy 1502 CTTGGTATGAGCTTACGCCCGCGGACACACAGTTAGGCTACGAGCATATACAAACACC 1561  
Db 4290 CTTGGTATGAGCTTACGCCCGCGGAGACTACAGTTAGGCTACGAGCGTATAGAACACC 4231  
Qy 1562 CGGGAATTCCCGTGTGCCAAGACCATCTTTGAATTTTGGGAGGCGCTTTTACGGGTCTCA 1621  
Db 4230 CGGGGCTTCCCGTGTGCCAAGACCATCTTTGAATTTTGGGAGGCGCTTTTACAGGCTCA 4171  
Qy 1622 CCACATAGACGCCCATCTTCTCTATCCAGACAAAGCAGAGTGGGGAACCTTTCCTATC 1681  
Db 4170 CTATATAGATGCCCATCTTCTATCCAGACAAAGCAGAGTGGGAGAACCTTTCCTTACC 4111  
Qy 1682 TGGTAGCTACCAAGCACCGTGTGCGTAGAGCTCAAGCCCTCCATGGGCCAAACACCTCTCTAT 1741  
Db 4110 TGGTAGCTACCAAGCACCGTGTGCGTAGAGCTCAAGCCCTCCATGGGCCAAACACCTCTCTAT 4051  
Qy 1742 AGATGTGGAAGTCTTTCGCTCAAGCCCAACCTCCATGGGCCAAACACCTCTCTAT 1801  
Db 4050 AGATGTGGAAGTCTTTCGCTCAAGCCCAACCTCCATGGGCCAAACACCTCTCTAT 3991  
Qy 1802 ATAGACTGGCGCTGTCCAGAAATGAAGTCACTGTGACGACCCAGTCAACAGTATATCA 1861  
Db 3990 ACAGACTGGCGCTGTTCAGAAATGAATCACTGTGACGACCCAGTCAACAGTATATCA 3931  
Qy 1862 TGACATGTATGCGCTGACCTGGAGTGTGTCAGAGTACCTGGGTGCTGTTGGCGGG 1921  
Db 3930 TGACATGTATGCGCTGACCTGGAGTGTGTCAGAGTACCTGGGTGCTGTTGGCGGG 3871  
Qy 1922 TTCTGCTGCTTTGGCGCGTATTCCTATCCACAGCTCCGTTGGTCTATAGTAGTAGGA 1981  
Db 3870 TCTGCTGCTTTGGCGCGTATTCCTATCCACAGCTCCGTTGGTCTATAGTAGTAGGA 3811  
Qy 1982 TTGCTTTGTCGGAAGCGGCAATCATACCCGACAGGGAAGTCTCTACCGGAGTTTCG 2041  
Db 3810 TCGTCTTTGTCGGAAGCGGCAATCATACCTGACAGGGAAGTCTCTACCGAGAGTTTCG 3751  
Qy 2042 ATGMAATGGAAGTCT 2059  
Db 3750 ATGAGATGGAAGTCT 3733

## RESULT 14

US-08-444-818-176  
; Sequence 176, Application US/08444818  
; Patent No. 6150087  
; GENERAL INFORMATION:  
; APPLICANT: Chien, William J.  
; APPLICANT: Rutter, William J.  
; TITLE OF INVENTION: NANBV Diagnostics and Vaccines  
; NUMBER OF SEQUENCES: 777  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Chiron Corporation  
; STREET: 4560 Horton Street  
; CITY: Emeryville  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 94608-2916  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/444,818  
; FILING DATE:  
; CLASSIFICATION: 424  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/08/403,590  
; FILING DATE: 14-MAR-1995  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Harbin, Alisa A.



Db 4536 AGCTGTCGATTTGGGCAATCAATGCCGTGGCTTACTACCGCGTCTTGACGTCCTCGTCA 4595  
Qy 1202 TCCGACCACTGTGTGCTGTCTGTGCAACTGACGCGCTCATGACCGGCTTTACCG 1261  
Db 4596 TCCGACCACTGTGTGTGTGTGCAACCGATGCTGTGCTGTGCAACCGCTTTACCG 4655  
Qy 1262 GCGACTTCGATTCGGTGATGACTGCAACGCTGTGTGCTGTGCAACCGCTTTACCG 1321  
Db 4656 GCGACTTCGATTCGGTGATGACTGCAACGCTGTGTGCTGTGCAACCGCTTTACCG 4715  
Qy 1322 TTGACCCCTACTTCAATTTGAGCAATCAAGCTTCCCGAGGATGCTGTCTCCCGTACTC 1381  
Db 4716 TTGACCCCTACTTCAATTTGAGCAATCAAGCTTCCCGAGGATGCTGTCTCCCGTACTC 4775  
Qy 1382 AAGCTCGGGGTAGGACTGGCAGAGGGAAGCAGGCACTTACAGATTTGTGSCACCGGGG 1441  
Db 4776 AAGCTCGGGGTAGGACTGGCAGAGGGAAGCAGGCACTTACAGATTTGTGSCACCGGGG 4835  
Qy 1442 AGCGTCCCTTCTGGCATGTTTGACTCGTCTGTCTCTCGAGTGTCTATGACGCGGTTGTG 1501  
Db 4836 AGCGCCCTCCGCAATGTTTCGACTCGTCTGTCTGTGAGTGTCTATGACGCGGTTGTG 4895  
Qy 1502 CTTGGTATGAGCTTACGCCCGCGAGACACAGTTAGGCTACAGACATATGAAACACC 1561  
Db 4896 CTTGGTATGAGCTTACGCCCGCGAGACTACAGTTAGGCTACAGACATATGAAACACC 4955  
Qy 1562 CGGACTTCCCGTGTGCAAGACCATCTTGAATTTTGGGAGGCGTCTTTACGGGCTCA 1621  
Db 4956 CGGACTTCCCGTGTGCAAGACCATCTTGAATTTTGGGAGGCGTCTTTACGGGCTCA 5015  
Qy 1622 CCACATAGACGCCACTTCTCTATCCAGACAAAGCAGAGTGGGGAACCTTCCCTATC 1681  
Db 5016 CTATATAGATGCCACTTCTCTATCCAGACAAAGCAGAGTGGGGAACCTTCCCTATC 5075  
Qy 1682 TGGTAGGCTACCAAGCACCGTGTGCGCTAGAGCTCAAGCCCTCTCCCGCTGTGGGACC 1741  
Db 5076 TGGTAGGCTACCAAGCACCGTGTGCGCTAGAGCTCAAGCCCTCTCCCGCTGTGGGACC 5135  
Qy 1742 AGATGTGAAGTGTGCTGATCCGCTCAAGCCCAAGCTTCAAGTGGGCAACCTCTGCTAT 1801  
Db 5136 AGATGTGAAGTGTGCTGATCCGCTCAAGCCCAAGCTTCAAGTGGGCAACCTCTGCTAT 5195  
Qy 1802 ATAGACTGGCGCTGTCCAGATGAAGTCAACCTGACGCAACCGAGTCAACCAAGTATATCA 1861  
Db 5196 ACAGACTGGCGCTGTTCAGATGAAGTCAACCTGACGCAACCGAGTCAACCAAGTATATCA 5255  
Qy 1862 TGACATGATATGTCGGCTGACCTGGAGTGTCTACGAGTACTGGGTGCTCGTTGGCGCG 1921  
Db 5256 TGACATGATATGTCGGCGGACCTGGAGTGTCTACGAGTACTGGGTGCTCGTTGGCGCG 5315  
Qy 1922 TTCTGGCTGCTTTGGCGCGGATTTGCTATCCACGCGTCCGCTGATAGTAGTAGGA 1981  
Db 5316 TCCTGGCTGCTTTGGCGCGGATTTGCTATCCACGCGTCCGCTGATAGTAGTAGGA 5375  
Qy 1982 TTCTGCTTGTCCGGAAGCCGCAATCATATCCGACAGGGAAGTCTCTACCGGAGTTCCG 2041  
Db 5376 TCCTGCTTGTCCGGAAGCCGCAATCATATCCGACAGGGAAGTCTCTACCGGAGTTCCG 5435  
Qy 2042 ATGAATGGAAGTGTCT 2059  
Db 5436 ATGAGATGGAAGTGTCT 5453

RESULT 15  
US-09-881-239-2  
; Sequence 2, Application US/09881239  
; Patent No. 6630298  
; GENERAL INFORMATION:  
; APPLICANT: CHIEN David Y.  
; APPLICANT: ARANGEL, Phillip  
; APPLICANT: TANDESKE, Laura  
; APPLICANT: GEORGE-NASCIEMENTO, Carlos  
; APPLICANT: COIT, Doris

; APPLICANT: MEDINA-SELBY, Angelica  
; TITLE OF INVENTION: HCV ANTIGEN/ANTIBODY COMBINATION ASSAY  
; FILE REFERENCE: 2302-16073 / PP16073.003  
; CURRENT APPLICATION NUMBER: US/09/881,239  
; CURRENT FILING DATE: 2001-06-14  
; NUMBER OF SEQ ID NOS: 8  
; SOFTWARE: Patent in Ver. 2.0  
; SEQ ID NO 2  
; LENGTH: 2058  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence:  
; OTHER INFORMATION: representative NS3/4a conformational antigen  
; NAME/KEY: CDS  
; LOCATION: (1)..(2058)  
US-09-881-239-2

Query Match 86.4%; Score 1781.2; DB 4; Length 2058;  
Best Local Similarity 91.6%; Pred. No. 0;  
Matches 1885; Conservative 0; Mismatches 173; Indels 0; Gaps 0;  
Qy 1 ATGGCCCTTATCAGCGCCTATGCCAGCAGACAGAGGGGCTTTGGGATGCATAATCAC 60  
Db 1 ATGGCCCTTATCAGCGCCTATGCCAGCAGACAGAGGGGCTTTGGGATGCATAATCAC 60  
Qy 61 AGCTTCAACCGCGCGGACAAACACAGGTGGAGGTGAGGTTCCAGATCGTGTCAACTGCT 120  
Db 61 AGCTTCACTGGCCGGACAAACACAGGTGGAGGTGAGGTTCCAGATCGTGTCAACTGCT 120  
Qy 121 GCCAGACTTTCTTGGCAACCTGCAATTAACGGGGTGTGTGGACTGTCTACCAATGAGCC 180  
Db 121 GCCCAAACTTCTTGGCAACCTGCAATTAACGGGGTGTGTGGACTGTCTACCAATGAGCC 180  
Qy 181 GGAACAGAGCAATTTCCGTACCTAAGGGTCTCTTATCCAGATGTACCAATGTGAC 240  
Db 181 GGAACAGAGCAATTTCCGTACCTAAGGGTCTCTTATCCAGATGTATACCAATGTAGAC 240  
Qy 241 CAAGACTCGTAGGCTGGCCGCTCCCAAGGTGCGGCTCAATTAACACACCATGCACTTGC 300  
Db 241 CAAGACTTTGTGGCTGGCCGCTCCCAAGGTGCGGCTCAATTAACACCATGCACTTGC 300  
Qy 301 GGCTCTCTCGACCTTTTACCTGTGTACAGGCAACCGGATGTCTATCTGTGCGCCGACGG 360  
Db 301 GGCTCTCTCGACCTTTTACCTGTGTACAGGCAACCGGATGTCTATCTGTGCGCCGCGG 360  
Qy 361 GGTGATGAGGCGGAGCTGTCTTCCCGCGGCTATCTCTTACTTGAAGGCTCTCTCG 420  
Db 361 GGTGATGAGGCGGAGCTGTCTTCCCGCGGCTATCTCTTACTTGAAGGCTCTCTCTCG 420  
Qy 421 GGAGGCGCTCTGTGTGCGCGGAGACATGCCGTAGGCAATTTCAAGCGCGGTATGC 480  
Db 421 GGAGGCTCCGTGTGTGCGCGGAGACATGCCGTAGGCAATTTCAAGCGCGGTATGC 480  
Qy 481 ACCCGTGGAGTGGCTAAGCGGTGGACTTTATCCCTGTGGAGAACCTTAGAGCAACCATG 540  
Db 481 ACCCGTGGAGTGGCTAAGCGGTGGACTTTATCCCTGTGGAGAACCTTAGAGCAACCATG 540  
Qy 541 AGGTCCCGGCTGTCTCAGACAACTCTCCCGACAGAGTGGCCCGAGCTACCAAGTG 600  
Db 541 AGGTCCCGGCTGTCTCAGACAACTCTCTCCCGACAGAGTGGCCCGAGCTACCAAGTG 600  
Qy 601 GCCCACTGCTATGCCCGGAGGTAAGAGCAACCAAGGTCCCGGCGCATACGCA 660  
Db 601 GCTCCTCTCATGCTCCCGAGGAGGCAACCAAGGTCCCGGCGCATATGCA 660  
Qy 661 GCTCAGGCTTACAGGTGTGTGTCTCAACCCCTCCCTGTCTGCAACAAATTTGGT 720  
Db 661 GCTCAGGCTTATAAGGTGTGTGTCTCAACCCCTCTGTCTGCAACAAATTTGGT 720  
Qy 721 GCTTACATGTCCAGGCGCATGGGATGATCTTAACTCAGGACTGGGTGAGGACAAAT 780  
Db 721 GCTTACATGTCCAGGCTCATGGGATGATCTTAACTCAGGACTGGGTGAGGACAAAT 780





**THIS PAGE BLANK (USPTO)**

C	1	46.8	2.3	509	9	CL252374	ZMMEBB060
C	2	44.8	2.2	935	9	CNS0065XK	Drosophill
C	3	44.4	2.2	834	4	BI956973	HVSMen000
C	4	43.4	2.1	666	5	BM950570	UI-M-EHOP
C	5	42.8	2.1	1863	7	CV068974	f2 new ch
C	6	42.6	2.1	833	4	BI959933	HVSMen002
C	7	42.2	2.0	871	6	CD437613	EL01N0503
C	8	41.8	2.0	564	2	BE291962	601085590
C	9	41.8	2.0	673	4	BI1155705	AG2004453
C	10	41.8	2.0	935	5	BQ876256	AGENCOURT
C	11	41.8	2.0	1020	6	BY711719	BM11719
C	12	41.8	2.0	1409	3	AK012576	Mus muscu
C	13	40.8	2.0	925	9	CNS00931P	Drosophill
C	14	40.4	2.0	533	4	HJ277484	BSJ277484
C	15	40.4	2.0	1749	9	CL948408	OSrFSB005
C	16	40.2	2.0	574	4	BJ208789	TJ208789
C	17	40.2	2.0	629	4	BQ306349	TaLr1148G
C	18	40.2	2.0	684	4	BI956133	HVSMen002
C	19	40	1.9	697	7	CN944784	011003AVB
C	20	39.8	1.9	624	6	CD338118	OV.109B20
C	21	39.8	1.9	896	4	BQ369273	602836743
C	22	39.6	1.9	424	5	BQ665971	HZ01IL20u
C	23	39.6	1.9	429	5	BQ665888	HZ01IH12u
C	24	39.6	1.9	430	5	BQ665954	HZ01IL02u



Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. Denatured mRNA was size fractionated on a 1% agarose gel. First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded cDNA was size selected according to mRNA size fraction, ligated with EcoR I adaptor, digested with Not I, and then cloned directionally into pXX-Asc vector. The library tag sequence located between the Not I site and the polyA tail, is CAGCCACGAC. This library was created for the University of Iowa Mouse Brain Molecular Anatomy Project (BMAP): 'Gene Discovery in the Developing Mouse Nervous System', supported by National Institute of Mental Health (NIMH), Henin Chin. Ph.D. program coordinator.

	Query Match	2.1%;	Score 43.4;	DB 5;	Length 656;	
	Best Local Similarity	53.3%;	Pred. No. 1.9;	Mismatches 76;	Indels 0;	Gaps 0;
	Matches	89;	Conservative 0;			
Qy	687	CAACCCCTCCGTTGTGCAACAATGGGCTTTGGTGTTACATGTCTCAAGGCCCATGGGAT	746			
Db	624	CAAGGCCACTGTGCTGGAGCCATGGGCTCTGCAGTGTCTCTGCCTCAGAGCCACAGGGCT	565			
Qy	747	TGATCTCTAATCAGACACTGGGGTGAGGCAATTACTTACTGCGAGCCGCCATCAGGTATTC	806			
Db	564	GGACACAGGGGAGTCTCTCGGGTAGGGCAAGTCCACAGGAGCGCTGGGATATGTGCC	505			
Qy	807	CACCTACGGCAAGTTCTCTTGGCCAGCCGGGTGTTTCAGGGGGTGC	851			
Db	504	CGGCTCTGGAGCAGCTGCAGCCGCCCGTGGGAGTTCAGCTGAGGC	460			

RESULT 5					
CV068974/c					
LOCUS	CV068974	1863 bp	mRNA	linear	EST 24-AUG-2004
DEFINITION	f2 New Chopped.fasta.Contig672 Preamplified custom cDNA library in pCwspOrt6.1 (ResGen, Invitrogen Inc.) Emiliana huxleyi cDNA, mRNA sequence.				
ACCESSION	CV068974				
VERSION	CV068974.1	GI:51532138			
KEYWORDS	EST.				
SOURCE	Emiliania huxleyi				
ORGANISM	Emiliania huxleyi				
REFERENCE	Eukaryota; Haptophyceae; Isochrysidales; Emiliania.				
AUTHORS	1 (bases 1 to 1863)				
TITLE	Wahlund,I.W., Zhang,X. and Read,B.A. Expressed Sequence Tag Profiles from Calcifying and Non-Calcifying Cultures of Emiliania huxleyi				
JOURNAL	Micropaleontology (2004) In press				

CONTACT: Betsy Read  
Department of Biological Sciences  
California State University San Marcos  
333 S. Twin Oaks Valley Road, San Marcos, CA 92096-0001, USA  
Tel: 760 750 4129  
Email: [bread@csusm.edu](mailto:bread@csusm.edu)  
Location/Qualifiers  
1. .i863  
/organism="Emiliana huxleyi"

```

ORIGIN
Query Match 2.1%; Score 42.8; DB 7; Length 1863;
Best Local Similarity 48.4%; Pred. No. 3.7;
Matches 119; Conservative 0; Mismatches 127; Indels 0; Gaps 0;

QY 1126 GACGAGCTCGCGCAAAACTGGTCGGTGTGGGCGTCAATGCGGTGCTTACTACGCGGC 1185

```

```

|||||
1523 GAGCCACAGCGCCATAGCGGACCCCTGCTCATAGCGCCCTCCCCCATTTGCC 1464
|||||
1186 CTTGATGTGTCCTATCCCGACCACTGTGTGAGTGTGTCGTCGTAAGTACAGCCCTC 1245
|||||
1463 GCGGGGTTTACCGGCTTCCCGGGCGCGCCGCGGAGGCGCTTGCCACCGCGCGCTTC 1404
|||||
1246 ATGACCGGCTTTACCGGCGACTTCGATTCCGTCATAGATGCAACAGTGTGTACCCAG 1305
|||||
1403 AAGCCACACCTTCCCGCGCGGCTGTGCGCGCTTGCCGCGCACCGCGCGCTTGTCTTT 1344
|||||
1306 ACAGTGTGACTTACGCTTACGCTTACCTTACCTTACCATTCAGACATACAGCTTCCCAAGGAT 1365
|||||
1343 GTGCGCGGCTTCCCGCCATCGCCACTTTCGCCACCGTCGCCATCGCGCACCGTCGGAT 1284
|||||
1366 GTGTGC 1371
|||||
1283 TCTGAC 1278

RESULT 6
BI959933
LOCUS
DEFINITION
HVSMen0022H07f Hordeum vulgare rachis EST library HVCDNA0015
(normal) Hordeum vulgare subsp. vulgare cDNA clone HVSMen0022H07f,
mRNA sequence.
BI959933
VERSION
KEYWORDS
SOURCE
ORGANISM
Hordeum vulgare subsp. vulgare
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Pooideae; Triticeae; Hordeum.
1 (bases 1 to 633)
Wing, R., Close, T.J., Klein, H., A., Wise, R., Chin, A., Begum, D.,
Frisch, D., Atkins, M., Yu, Y., Henry, D., Palmer, M., Rambo, T.,
Simmons, J., Oates, R., and Main, D.
Development of a genetically and physically anchored EST resource
for barley genomics: Morex rachis cDNA library
Unpublished (2001)
Contact: Wing RA
Clemson University Genomics Institute
Clemson University
100 Jordan Hall, Clemson, SC 29634, USA
Tel: 864 656 7288
Fax: 864 656 4293
Email: rwing@clemson.edu
Total hg bases = 513
Seq primer: AATTAAACCTTCACTAAAGGG
High quality sequence stop: 595.
Location/Qualifiers
1..633
/organism="Hordeum vulgare subsp. vulgare"
/mol_type="mRNA"
/cultivar="Morex"
/sub_species="vulgare"
/db_xref="taxon:112509"
/clone="HVSMen0022H07f"
/tissue_type="Rachis"
/lab_host="TJC121"
/clone_lib="Hordeum vulgare rachis EST library HVCDNA0015
(normal)"
/note="vector: pBluescript SK(-); Site 1: EcoRI; Site 2:
XhoI; Plants were grown at Washington State University,
Pullman, WA in a greenhouse, the rachis were excised and
frozen in liquid nitrogen (Klein, H. lab). In the TJ Close
lab at the University of California, Riverside total RNA
was prepared. poly(A) was purified, one primary
unamplified cDNA library was made, and 1 million pfu were
in vivo excised to give pBluescript SK(-) cDNA phagemids
(Chin). Phagemids were plated and picked at the Clemson
University Genomics Institute (CUGI) (Begum, Palmer,

```

```

Frisch, Atkins and Wing). Plasmid DNA preparations, DNA
sequencing and sequence analysis were performed at CUGI
(Wing, Yu, Frisch, Henry, Simmons, Rambo, Main). The
sequence has been trimmed to remove vector sequence and
contains a minimum of 100 bases of phred value 20 or
above. For more details on library preparation and
sequence analysis see
http://www.genome.clemson.edu/projects/barley. To order
this clone see http://www.genome.clemson.edu/orders/Also
see Close TJ, Wing R, Klein, H., A., Wise R (2001)
Genetically and physically anchored EST resources for
barley genomics. Barley Genetics Newsletter 31:23-30.
(http://wheat.pw.usda.gov/ggpages/bgn/31/cover.html)."

ORIGIN
Query Match 2.1%; Score 42.6; DB 4; Length 633;
Best Local Similarity 51.3%; Pred. No. 3.1;
Matches 99; Conservative 0; Mismatches 94; Indels 0; Gaps 0;

QY 1096 CTCATCTTCTGCCACTCAAAAGAGAGTGCAGACGAGCTCGCGCAAACTGGTCGGTTG 1155
|||||
405 CTCCTCACTGTCCAGCGGACAGACTCGCGGGCCCCAACGGGACGGTTCGGCGG 464
|||||
1156 GCGCTCAATGCGGTGCTTACTACCGCGCTTGTATGTGTCTCGTATCCGACCACTGGT 1215
|||||
465 AGCATCAACACGCTCTCTGCTGCTCCCTCCACCGTCTCCATCTCCAGGCACACTAC 524
|||||
1216 GAGTGTGTGTGTCGCACTGACCGCTTCATGACCGGCTTTACCGCGGACATTCGATTG 1275
|||||
525 CAGGGCGATGCGCGCGCCACCGCGGCTTCCACCGGACTTTCCCGCCCAACCGCGG 584
|||||
1276 GTGATAGACTGCA 1288
|||||
585 CAGTTCGACTACA 597
|||||

RESULT 7
CD437613
LOCUS
DEFINITION
EL01N0503A05.b Endosperm_5 Zea mays cDNA, mRNA sequence.
CD437613
VERSION
KEYWORDS
SOURCE
EST.
ORGANISM
Zea mays
Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
1 (bases 1 to 871)
Lai, J., Dey, N., Kim, C.S., Bharti, A.K., Rudd, S., Mayer, K.F.,
Larkins, B., Becraft, P., and Messing, J.
Characterization of the maize endosperm transcriptome and its
comparison to the rice genome
Genome Res. 14 (10), 1932-1937 (2004)
Contact: Lai, Jinsheng
Dr. Joachim Messing's lab
Waksman Institute, Rutgers University
190 Frelinghuysen Rd., Piscataway, NJ 08854, USA
Tel: 732-445-3801
Fax: 732-445-5735
Email: jlai@waksman.rutgers.edu
Seq primer: T3.
Location/Qualifiers
1..871
/organism="Zea mays"
/mol_type="mRNA"
/cultivar="W22"
/db_xref="taxon:4577"
/tissue_type="Endosperm of 7-23DAP"
/clone_lib="Endosperm 5"
/note="Vector: pBluescript SK-; Site 1: EcoRI; Site 2:
XhoI"

ORIGIN

```



[illegible][illegible]

```

RESULT 10
BQ876256/c
LOCUS
DEFINITION BQ876256 935 bp mRNA linear EST 16-AUG-2002
AGENCOURT 8743841 NIH_MGC_129 Mus musculus cDNA clone IMAGE:6389941
5', mRNA sequence.
ACCESSION BQ876256
KEYWORDS BQ876256.1 GI:22268262
SOURCE EST.
ORGANISM Mus musculus (house mouse)
Mammalia; Euthera; Chordata; Vertebrata; Euteleostomi;
Mammalia; Euthera; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 935)
NIH-MGC http://mgc.nci.nih.gov/
NATIONAL INSTITUTES OF HEALTH, Mammalian Gene Collection (MGC)
UNPUBLISHED (1999)
CONTACT: Robert Strausberg, Ph.D.
EMAIL: cga@rs-rcmail.nih.gov
Tissue Procurement: Susan L. Sullivan, PhD.
CDNA Library Preparation: ResGen, Invitrogen Corp
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L14M13875 row: m column: 14
High quality sequence stop: 761.
FEATURES
source
Location/Qualifiers
1..935
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="IMAGE:6389941"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_129"
/note="Organ: olfactory epithelium; Vector: pCMV-SPORT6.1;
Site1: EcoRV; Site2: NotI; Cloned unidirectionally.
Primer: Oligo dt. Average insert size 2.2 kb. Constructed
by ResGen, Invitrogen Corp. Note: this is a NIH_MGC
Library."
ORIGIN
Query Match 2.0%; Score 41.8; DB 5; Length 935;
Best Local Similarity 53.3%; Pred. No. 5.8;
Matches 88; Conservative 0; Mismatches 77; Indels 0; Gaps 0;
687 CAACCCCTCGTGTGCTGCAACATGGGCTTTGGTGTTCATCTCCAGGCCCATGGAT 746
|||||
548 CAAGGCCACCTGTGCTGGAGCCATGGGCTCTGCAGTCTCTCGTCTCAGAGCCACAGGCT 489
|||||
747 TGATCTTAACATCAGGACTGGGTGAGGACAATTAATCTGCGAGCCCGATCATGTTTC 806
|||||
488 GGACACAAGGGGAGCTCTCTCGGTAGGGCAAGGTCACAGGAGCGCTGGGCTAGTGCC 429
|||||
807 CACTACGGCAAGTCTCTTCCGACGCGGCTGTCAGGGGTCG 851
|||||
428 CGGCTCTGGAGCAGCTCGACGCCGCGGTGGAGTTTCAGCTGAGCG 384
|||||
RESULT 11
BY711719/c
LOCUS
DEFINITION BY711719 RIKEN full-length enriched, 11 days embryo whole body Mus
musculus cDNA clone 2700087H15 5', mRNA sequence.
ACCESSION BY711719
KEYWORDS BY711719.1 GI:27122976
SOURCE EST.
ORGANISM Mus musculus (house mouse)
Mammalia; Euthera; Chordata; Vertebrata; Euteleostomi;
Mammalia; Euthera; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 1020)

```

## AUTHORS

Okazaki, Y., Furuno, M., Kasukawa, T., Adachi, J., Bono, H., Kondo, S., Nikaide, I., Osato, N., Saito, R., Suzuki, H., Yamanaka, I., Kiyosawa, H., Yagi, K., Tomaru, Y., Hasegawa, Y., Nogami, A., Schonbach, C., Gojobori, T., Baldarelli, R., Hill, D. P., Bult, C., Hume, D. A., Quackenbush, J., Schriml, L. M., Kanapin, A., Matsuda, H., Batalov, S., Beisel, K. W., Blake, J. A., Bradt, D., Brusic, V., Chothia, C., Corbani, L. E., Cousins, S., Dalla, E., Dragani, T. A., Fletcher, C. F., Forrest, A., Frazer, K. S., Gasterland, T., Gariboldi, M., Gissi, C., Godzik, A., Gough, J., Grimmond, S., Gustincich, S., Hirokawa, N., Jackson, I. J., Jarvis, E. D., Kanai, A., Kawaji, H., Kawasawa, Y., Kedzierski, R. M., King, B. L., Konagaya, A., Kurochkin, I. V., Lee, Y., Lenhard, B., Lyons, P. A., Maglott, D. R., Maltais, L., Marchionni, L., McKenzie, L., Miki, H., Nagashima, T., Numata, K., Okido, T., Pavan, W. J., Perlea, G., Pesole, G., Petrovsky, N., Pillai, R., Pontius, J. U., Qi, D., Ramachandran, S., Ravasi, T., Reed, J. C., Reed, D. J., Reid, J., Ring, B. Z., Ringwald, M., Sandelin, A., Schneider, C., Sempile, C. A., Setou, M., Shimada, K., Sultana, R., Takenaka, Y., Taylor, M. S., Teasdale, R. D., Tomita, M., Verardo, R., Wagner, L., Wahlestedt, C., Wang, Y., Watanabe, Y., Wells, C., Wilming, L. G., Wynshaw-Boris, A., Yanagisawa, M., Yang, I., Yang, L., Yuan, Z., Zavolan, M., Zhu, Y., Zimmer, A., Carninci, P., Hayatsu, N., Hirozane-Kitahara, T., Konno, H., Nakamura, M., Sakazume, N., Sato, K., Shiraki, T., Waki, K., Kawai, J., Aizawa, K., Arakawa, T., Fukuda, S., Hara, A., Hashizume, W., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Miyazaki, A., Sakai, K., Sasaki, D., Shibata, K., Shinagawa, A., Yasunishi, A., Yoshino, M., Waterston, R., Lander, E. S., Rogers, J., Birney, E. and Hayashizaki, Y.

Analysis of the mouse transcriptome based on functional annotation

of 60,770 full-length cDNAs

Nature 420, 563-573 (2002)

22354683

12466851

COMMENT

Contact: Yoshihide Hayashizaki

Laboratory for Genome Exploration Research Group, RIKEN Genomic

Sciences Center (GSC), Yokohama Institute

The Institute of Physical and Chemical Research (RIKEN)

1-7-2 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan

Tel: 81-45-503-9222

Fax: 81-45-503-9216

Email: genome-res@sc.riken.jp, URL: http://genome.gsc.riken.jp/

Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Carninci, P.,

Fukuda, S., Hashizume, W., Hayashida, K., Hirozane, T., Hori, F.,

Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kawai, J., Kojima, Y.,

Kondo, S., Konno, H., Koya, S., Miyazaki, A., Murata, M., Nakamura, M.,

Nomura, K., Nomura, R., Ohno, M., Ohsato, N., Saito, R., Sakazume, N.,

Sano, H., Sasaki, D., Sato, K., Shibata, K., Shiraki, T., Tagami, M.,

Takeda, Y., Waki, K., Watanabe, A., Muramatsu, M. and Hayashizaki, Y.

Direct Submission

Computational Analysis of Full-Length Mouse cDNAs Compared with

Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)

Normalization and subtraction of cap-trapper-selected cDNAs to

prepare full-length cDNA libraries for rapid discovery of new

genes. Genome Res. 10 (10), 1617-1630 (2000)

RIKEN integrated sequence analysis (RISA) system-384-format

sequencing pipeline with 384 multicapillary sequencer. Genome Res.

10 (11), 1757-1771 (2000)

Computer-based methods for the mouse full-length cDNA

encyclopedia: real-time sequence clustering for construction of a

nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)

cDNA library was prepared and sequenced in Mouse Genome

Encyclopedia Project of Genome Exploration Research Group in Riken

Genomic Sciences Center and Genome Science Laboratory in Riken.

Division of Experimental Animal Research in Riken contributed to

prepare mouse tissues.

Please visit our web site (<http://genome.gsc.riken.go.jp>) for

further details.

Location/Qualifiers

1..1020

/organism="Mus musculus"

/mol\_type="mRNA"

/strain="C57BL/6J"

/db\_xref="taxon:10090"

/clone="2700087H15"

## FEATURES

source



```
QY 747 TGATCTTAACATCAGGACTGGGGTAGGACAAATTACTTGGCAGCGCCGATCAGTATTC 806
Db 292 GGACACAGGGGAGCTCTCGGGTAGGCAAGGTCCACAGGAGCCTCGGGGTAGTGCC 233
QY 807 CACTACGGCAAGTTCCTTCCGACAGCGCGGGTTCAGGGGGTGC 851
Db 232 CGGCTCTGGAGCAGCTGCAGCGCGCGGTGGAGATTCAGCTGAGGC 188

RESULT 13
CNS0091P/c 925 bp DNA linear GSS 03-JUN-1999
LOCUS Drosophila melanogaster genome survey sequence TET3 end of BAC #
DEFINITION BACR19D16 of RPCI-98 library from Drosophila melanogaster (fruit
fly), genomic survey sequence.
ACCESSION AL053013.1 GI:4934461
VERSION AL053013.1
KEYWORDS GSS.
SOURCE Drosophila melanogaster (fruit fly)
ORGANISM Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
REFERENCE 1 (bases 1 to 925)
AUTHORS Direct Submission
TITLE Genoscope.
JOURNAL Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
COMMENT Determination of this BAC-end sequence was carried out as part of a
collaboration with the Berkeley Drosophila Genome Project (BDGP).
The BDGP is constructing a physical map of the Drosophila
melanogaster genome using these BACs. For further information
please see http://www.fruitfly.org/TheBDGP/Drosophila
melanogaster library was prepared by Kazutoyo Osoegawa and
Aaron Mamoser in Pieter de Jong's laboratory in the Department of
Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,
NY. The library is named RPCI-98 and was constructed by partial
EcoRI digestion of Drosophila DNA provided by the BDGP's
isogenic strain Y2; cn bw sp, the same strain used for the BDGP's
PI and EST libraries. A more detailed description of the library
and how to order individual BAC clones, the entire library, or
filters for hybridization from the BACPAC Resource Center can be
found at http://bacpac.med.buffalo.edu/drosophila\_bac.htm.

FEATURES             Location/Qualifiers
     source
     1..925
         /organism="Drosophila melanogaster"
         /mol_type="genomic DNA"
         /db_xref="taxon:7227"
         /clone="BACR19D16"
         /clone_lib="RPCI-98"
         /note="end : TET3"

ORIGIN
Query Match      2.0%; Score 40.8; DB 9; Length 925;
Best Local Similarity 12.9%; Pred. No. 11;
Matches 47; Conservative 159; Mismatches 158; Indels 0; Gaps 0;

QY 346 CTGTGCGCCGAGGTGATGCGAGGGGAGCGCTGTTTCCGCCCGGCGCTATCTCTTAC 405
Db 916 SCSSBSCSSSMETSSSSBSCSSSSBSSSTSSSMSSSSBSSSSSSSSSSSSSSSS 857
QY 406 TTGAAAGGCTCTCGGAGGCCCTCTGCTGCCCGCAGACATGCCGTAGGCATATTC 465
Db 856 KCMASSCGCGGMBACMCSSSSCCGASARGVKVRASGGAGKRGGGSGSASHS 797
QY 466 AGAGCCGCGGTATGCAACCCGTGGGTGAGCTTCATCCCGGTAGAGAGC 525
Db 796 SSACBSSSSSCASCSWSSSSSSASRSRSGGAGSGASSRSSSSSSSASAGSVSS 737
QY 526 TTAGAGCAACCATGAGGTCCCGGTGTTCTGACAACTCTCCCAACGACAGGTGCC 595
Db 736 ASSSSSSSCSSSVSSVSSMCSBSSSSSSASASSSSSSSSSSSSSSSSSSSSS 677
```

```
QY 586 CAGAGCTACCAAGTGGCCACCTGCATGCTCCACCGCAGCGGTAAAGACCAAGGTC 645
Db 676 SAMSARSSSSSSSSSSSSMSASSSSSSSSSSSSSSSSSSSSSSSACGBSMSSGGSG 617
QY 646 CCGGCGGCATACGACGACTCAGGCTCAAGGTGCTGTCTCAACCCCTCGTTGTGCA 705
Db 616 SVSASSGMSVSSVSSGGRSSGSGGGVGGSSGSSSSSSSSGSGSVCSGSCMCRSC 557
QY 706 ACAA 709
Db 556 SSAA 553

RESULT 14
BJ277484/c 533 bp mRNA linear EST 09-APR-2002
LOCUS BJ277484 Y. Ogiwara unpublished cDNA library, Wh_r Triticum
DEFINITION aestivum cDNA clone whrl107 5', mRNA sequence.
ACCESSION BJ277484.1 GI:20100893
VERSION BJ277484
KEYWORDS EST.
SOURCE Triticum aestivum (bread wheat)
ORGANISM Triticum aestivum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Pooidae; Triticeae; Triticum.
REFERENCE 1 (bases 1 to 533)
AUTHORS Ogiwara, Y. and Murai, K.
TITLE Expressed genes in Triticum aestivum
JOURNAL Unpublished (2002)
COMMENT Contact: Tadasu Shin-i
Center For Genetic Resource Information
National Institute of Genetics
1111 Yata, Mishima, Shizuoka 411-8540, Japan
Tel: 81-559-81-6856
Fax: 81-559-81-6855
Email: tshini@genes.nig.ac.jp.

FEATURES             Location/Qualifiers
     source
     1..533
         /organism="Triticum aestivum"
         /mol_type="mRNA"
         /cultivar="Chinese Spring"
         /db_xref="taxon:4565"
         /clone="whrl107"
         /tissue_type="root"
         /dev_stage="freekes' scale 1"
         /clone_lib="Y. Ogiwara unpublished cDNA library, Wh_r"

ORIGIN
Query Match      2.0%; Score 40.4; DB 4; Length 533;
Best Local Similarity 58.2%; Pred. No. 12;
Matches 71; Conservative 0; Mismatches 51; Indels 0; Gaps 0;

QY 1124 GGCACGAGTCCGCGCAAACTGGTCGCTGGCGCTCAATGCGGTGCTTACTACCGG 1183
Db 452 GCGGCGAATCGCGAAGAACTCGTCGGCGCGGTGGCATTCGCTGGCGCATCAGTAG 393
QY 1184 GCTTGATGTGTCGTCATCCGACAGTGGTGAAGTTCGTCGTGGCAACTGACGCC 1243
Db 392 GGCCTGTGTCTCCGTCGAGGAGCAGCGCGGCGTCTGGTGAGAGGAGGCCCTGCTCGCC 333
QY 1244 TC 1245
Db 332 GC 331

RESULT 15
CL948408 1749 bp DNA linear GSS 21-SEP-2004
LOCUS OeIFSB005002 Oryza sativa Express Library Oryza sativa (indica
DEFINITION cultivar-group) genomic, genomic survey sequence.
ACCESSION CL948408
```

Search completed: February 27, 2005, 22:36:51  
Job time : 6109 secs

**THIS PAGE BLANK (USPTO)**



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: February 27, 2005, 01:27:56 ; Search time 6776 Seconds  
(without alignments)  
4905.590 Million cell updates/sec

Title: US-09-930-591-2  
Perfect score: 3618  
Sequence: 1 MAPTAYAQTRGLGCIIT.....PALIPREVLYRFEDEMEEC 686

Scoring table:  
BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODEL=frame+ p2n.model -DEV=xlp  
-Q=/cgn2.1/USPTO\_spool\_p/US09930591/runat\_25022005\_143612\_25675/app\_query.fasta\_1.839  
-DB=GenEmbl -QWTF=fastcap -SUFFIX=rgc -MINWATCH=0.1 -LOOPCL=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45  
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=200000000  
-USER=US09930591.cGCGN 1.1 5600 @runat\_25022005\_143612\_25675 -NCPU=6 -ICPU=3  
-NO\_MMAP -LARGQUERY -NEG\_SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : GenEmbl.\*  
1: gb\_ba.\*  
2: gb\_hcg.\*  
3: gb\_in.\*  
4: gb\_on.\*  
5: gb\_ov.\*  
6: gb\_pat.\*  
7: gb\_ph.\*  
8: gb\_pl.\*  
9: gb\_pr.\*  
10: gb\_ro.\*  
11: gb\_sts.\*  
12: gb\_sv.\*  
13: gb\_un.\*  
14: gb\_vi.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Length	ID	Description
1	3618	100.0	2061	CQ826998
2	3618	100.0	2061	AX441176
3	3618	100.0	2061	AX467113
4	3618	100.0	2078	CQ827032

5	3579	98.9	8791	14	AY615798	AY615798 Hepatitis I05434 Sequence 48
6	3574	98.8	5360	6	I06434	I09328 Sequence 8
7	3574	98.8	5360	6	I09328	I06440 Sequence 54
8	3574	98.8	6785	6	I09329	I09329 Sequence 10
9	3574	98.8	6785	6	AR118696	AR118696 Sequence
10	3574	98.8	7310	6	AR118696	I09331 Sequence 15
11	3574	98.8	7310	6	I09331	M32084 Hepatitis C
12	3574	98.8	9185	6	I08294	I08294 Sequence 1
13	3574	98.8	9185	6	BD091382	BD091382 HCV culti
14	3574	98.8	9185	6	BD091382	AR166930 Sequence
15	3574	98.8	9379	6	AR166930	AR301300 Sequence
16	3574	98.8	9379	6	AR301300	E66593 Hepatitis C
17	3574	98.8	9401	6	AR176483	AR176483 Sequence
18	3574	98.8	9401	6	E66593	I71894 Sequence 9
19	3574	98.8	9401	6	I71894	I81885 Sequence 9
20	3574	98.8	9401	6	I81885	BD080334 Hepatitis
21	3574	98.8	9401	6	BD080334	M62331 Hepatitis C
22	3574	98.8	9401	14	HPCFLYPRE	AF387805 Synthetic
23	3574	98.8	9609	12	AF387805	AF387806 Synthetic
24	3574	98.8	9646	12	AF387806	AF387807 Synthetic
25	3574	98.8	9693	12	AF387807	AR118692 Sequence
26	3571	98.7	6785	6	AR118692	AR118703 Sequence
27	3570	98.7	8316	6	AR118703	AR118722 Sequence
28	3570	98.7	9185	6	AR118722	AR118723 Sequence
29	3570	98.7	9185	6	AR118723	AF271632 Hepatitis
30	3570	98.7	9618	14	AF271632	AX164584 Sequence
31	3568	98.6	6299	6	AX164584	AR404933 Sequence
32	3565	98.5	2058	6	AR404933	AR408362 Sequence
33	3565	98.5	2058	6	AR408362	AX395309 Sequence
34	3565	98.5	2058	6	AX395309	AX454818 Sequence
35	3565	98.5	2058	6	AX454818	AR118728 Sequence
36	3565	98.5	8987	6	AR118728	AF387808 Synthetic
37	3565	98.5	9609	12	AF387808	AR118686 Sequence
38	3551	98.1	9646	6	AR118686	AR118828 Sequence
39	3551	98.1	9646	6	AR118828	BD069982 Functiona
40	3551	98.1	9646	14	AF009606	AF009606 Hepatitis
41	3551	98.1	12980	6	AR110831	AR110831 Sequence
42	3551	98.1	12980	6	BD069985	BD069985 Functiona
43	3551	98.1	9395	14	AF511950	AF511950 Hepatitis
44	3550	98.1	9502	6	E08263	E08263 GRNA of Hep
45	3550	98.1				

ALIGNMENTS

RESULT 1  
CQ826998  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
FEATURES  
ORIGIN

CQ826998  
Sequence 1 from Patent WO2004048402.  
CQ826998  
CQ826998.1 GI:49455655  
synthetic construct  
synthetic construct  
other sequences; artificial sequences.  
1  
Sallberg, M.  
A hepatitis c virus codon optimized non-structural ns3/4a fusion  
Gene  
Patent: WO 2004048402-A 1 10-JUN-2004;  
TRIPEP AB (SE)  
Location/Qualifiers  
1..2061  
/organism="synthetic construct"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:32630"  
/note="Hepatitis C virus NS3/4A coding region"

CQ826998  
Sequence 1 from Patent WO2004048402.  
CQ826998  
CQ826998.1 GI:49455655  
synthetic construct  
synthetic construct  
other sequences; artificial sequences.  
1  
Sallberg, M.  
A hepatitis c virus codon optimized non-structural ns3/4a fusion  
Gene  
Patent: WO 2004048402-A 1 10-JUN-2004;  
TRIPEP AB (SE)  
Location/Qualifiers  
1..2061  
/organism="synthetic construct"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:32630"  
/note="Hepatitis C virus NS3/4A coding region"

ORIGIN

Alignment Scores:  
Pred. No.: 7.16e-203 Length: 2061  
Score: 3618.00 Matches: 686  
Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00%		Mismatches: 0
Query Match: 100.00%		Indels: 0
DB: 6		Gaps: 0
US-09-930-591-2 (1-686) x CQ826998 (1-2061)		
QY	1 MetAlaProIleThrAlaValAlaGlnThrArgGlyLeuLeuGlyCysIleIleThr 20	341 AlaLeuSerThrThrGlyGluIleProPheTyrGlyLeuAlaIleProLeuGluAlaIle 360
DB	1 ATGGCGCTTATCACGGCTATGCCAGCAGACAAAGGGCCCTTTGGATGCATAAATCAC 60	1021 GCTCTGTCACATACCGAGAGATCCCTTTATGGCAAGGCTATTCCCTTGAAGCAATT 1080
QY	21 SerLeuThrGlyArgAspLysAsnGlnValGluGlyGluValGlnIleValSerThrAla 40	361 LysGlyGlyArgHisLeuIlePheCysHisSerLysLysLysCysAspGluLeuAlaAla 380
DB	61 AGCTTGACCGCGCGGACAAAACAGGTGGAGGTTGAGGTTCCAGATCGTGCACTGCT 120	1081 AAGGGGGGAGACATCTCATCTCTGCCACTCAAGAAGAAGTGCAGACGAGCTCGCGCA 1140
QY	41 AlaGlnThrPheLeuAlaThrCysIleAsnGlyValCysTrpThrValTyrHisGlyAla 60	381 LysLeuValAlaLeuGlyValAsnAlaValAlaTyrTyrArgGlyLeuAspValSerVal 400
DB	121 GCCAGACTTCTTGGCACTGCATTAACGGGGTGTGTGGACTGTCTACCATGGAGCC 180	1141 AAATCGTTCGGTGGCGCTCAATGCCGTGGCTTACTACCGGGCTTCAATGTGTCCGTC 1200
QY	61 GlyThrArgThrIleAlaSerProLysGlyProValIleGlnMetTyrThrAsnValAsp 80	401 IleProThrSerGlyAspValValValAlaAlaThrAspAlaLeuMetThrGlyPheThr 420
DB	181 GGAACAAGGACCATTCGGTCACCTAAGGTCCTGTTATCCAGATGTACACCAATGTGGAC 240	1201 ATCCCGACCAAGTGTGACGTTGTGCTGGCAACTACGCCCTCATGCCGCTTTTACC 1260
QY	81 GlnAspLeuValGlyTrpProAlaProGlnGlyAlaArgSerLeuThrProCysThrCys 100	421 GlyAspPheAspSerValIleAspCysAsnThrCysValThrGlnThrValAspPheSer 440
DB	241 CAAGACCTCGTAGGCTGGCCGCTCCCAAGGTGCCGCTCATTAACACCATGCACTTGC 300	1261 GGCAGCTTCGATTCGGTGATAGACTGCAACACGTTGTCCACCCAGACAGTCGACTTCAGC 1320
QY	101 GlySerSerAspLeuTyrLeuValThrArgHisAlaAspValIleProValArgArgArg 120	441 LeuAspProThrPheThrIleGluThrIleThrLeuProGlnAspAlaValSerArgThr 460
DB	301 GGCTCTCGGACCTTTACCTGGTGCAGAGGACGCGGATGTCATTTCTGTGGCCGACGG 360	1321 CTGACCCCTACCTTACCATTCAGCAATCAGCTTCCCAGGATGCTGTCTCCGCTACT 1380
QY	121 GlyAspGlyArgGlySerLeuLeuSerProArgProIleSerTyrLeuLysGlySerSer 140	461 GlnArgArgGlyArgThrGlyArgGlyLysProGlyIleTyrArgPheValAlaProGly 480
DB	361 GGTGTGGCAGGGCAGCGCTCTTTCGCCCGGCTATCTTACTTGAAGGCTCCTCG 420	1381 CAACTCGGGGTAGGACTGGCAGAGGAAGCCAGGCATCTACAGATTTGTGGCACCAGGG 1440
QY	141 GlyGlyProLeuLeuCysProAlaGlyHisAlaValGlyIlePheArgAlaAlaValCys 160	481 GluArgProSerGlyMetPheAspSerValLeuCysGluCysTyrAspAlaGlyCys 500
DB	421 GGAGGCGCTCTGTGTGCCCGCAGGACATGCGGTAGGCATATTCAGAGCCGGGTATGC 480	1441 GAGCGTCTTCTGGCATGTTTGAATCGTCTGCTCTGCGAGTGTATGACGCGGTTGT 1500
QY	161 ThrArgGlyValAlaLysAlaValAspPheIleProValGluSerLeuGluThrThrMet 180	501 AlaTrpTrpGluLeuThrProAlaGluThrThrValArgLeuArgAlaTyrMetAsnThr 520
DB	481 ACCCGTGGAGTGGCTAAGCGGTGGACTTCCATCCCGTAGAGCTTAGAGACACCATG 540	1501 GCTTGTATGAGCTTACGCCCGCCGAGACCAGATTAGGCTACGACATACATGAACACC 1560
QY	181 ArgSerProValPheSerAspAsnSerSerProAlaValProGlnSerTyrGlnVal 200	521 ProGlyLeuProValCysGlnAspHisLeuGluPheTrpGluGlyValPheThrGlyLeu 540
DB	541 AGTCCCCCGGTGTTCTCAGACAACCTCTCTCCCACAGCAGTGCCTCCAGAGCTACCAAGTG 600	1561 CCGGGACTTCCCGTGTGCCAAGACCATCTTGAATTTGGAGGGCGCTTTTACGGGTCTC 1620
QY	201 AlaHisLeuHisAlaProThrGlySerGlyLysSerThrLysValProAlaAlaTyrAla 220	541 ThrHisIleAspAlaHisPheLeuSerGlnThrLysGlnSerGlyGluAsnLeuProTyr 560
DB	601 GCCACCTGCATGCTCCACCGGCGAGCGGTAAAGACACCAAGTCCCGCGCATACGCA 660	1621 ACCCATAGACGCCCATCTTCTATCCAGACAAAGCAGAGTGGGAAAACCTTCCCTAT 1680
QY	221 AlaGlnGlyTyrLysValLeuValLeuAsnProSerValAlaAlaThrMetGlyPheGly 240	561 LeuValAlaTyrGlnAlaThrValCysAlaArgAlaGlnAlaProProSerTrpAsp 580
DB	661 GCTCAGGGCTACAAAGGTGGTGGTCTCAACCCCTCGGTCTGCTGCAACAATGGGCTTTGGT 720	1681 CTGGTAGCGGTACCAAGCCCGGTGCGGTAGAGCTCAAGCCCTCCCCCGTCTGGGAC 1740
QY	241 AlaTyrMetSerLysAlaHisGlyIleAspProAsnIleArgThrGlyValArgThrIle 260	581 GlnMetTrpLysCysLeuIleArgLeuLysProThrLeuHisGlyProThrProLeuLeu 600
DB	721 GCTTACATGTCGAAGGCCCATGGATTGATCTTAACATCAGACTGGGGTGAGACAAATT 780	1741 CAGATGTGAAGTGTCTGATCCGTCTCAAGCCACCCTCCATGGGCCAACACCTCTGCTA 1800
QY	261 ThrThrGlySerProIleThrTyrSerThrTyrGlyLysPheLeuAlaAspGlyCys 280	601 TyrArgLeuGlyAlaValGlnAsnGluValThrLeuThrHisProValThrLysTrpIle 620
DB	781 ACTACTGGACGCCGATCAGTATATCCACCTAGCGCAAGTTCCTTGGCCGCGGGGTGT 840	1801 TATAGACTGGCGCTCTCCAGAATGAAGTCACTCCAGCAGCCAGTCAGTCAAGTAGTATC 1860
QY	281 SerGlyGlyAlaTyrAspIleIleCysAspGluCysHisSerThrAspAlaThrSer 300	621 MetThrCysMetSerAlaAspLeuGluValThrSerThrTrpValLeuValGlyGly 640
DB	841 TCAGGGGGTCTTATGACATAATAATTGTGAGAGGTGCCACTCCAGGATGCAATCC 900	1861 ATGACATGATGTGCGCTACCTGGAGGTGCTCAGAGTACCTGGGTCTCGTTGGCGGC 1920
QY	301 IleLeuGlyIleGlyThrValLeuAspGlnAlaGluThrAlaGlyAlaArgLeuThrVal 320	641 ValLeuAlaAlaLeuAlaAlaTyrCysLeuSerThrGlyCysValValIleValGlyArg 660
DB	901 ATCTTGGCATTTGGCACTGCTTGTGACCAAGCAGAGACCGCGGGGCGAGACTGACTGTG 960	1921 GTTCTGGCTGCTTTGGCCGCGTATTGCCCTATCCACAGGCTGCGTGTGCTAGTAGTAGG 1980
QY	321 LeuAlaThrAlaThrProGlySerValThrValProHisProAsnIleGluGluVal 340	661 IleValLeuSerGlyLysProAlaIleIleProAspArgGluValLeuTyrArgGluPhe 680
DB	961 CTGCCACCGCTACCCCTCGGGCTCCGTCACCTGTGCCCATCTCTAACATCGAGGAGGTT 1020	1981 ATTGTCTTGTCCGAAAAGCCGCAATCATACCCGACAGGGAAGTCTCTCTACCGGAGTTC 2040
QY		681 AspGluMetGluGluCys 686
DB		2041 GATGAAATGGAAGAGTGC 2058

```
RESULT 2
AX441176
LOCUS AX441176 2061 bp DNA linear PAT 28-JUN-2002
DEFINITION Sequence 16 from Patent WO0213855.
ACCESSION AX441176
VERSION AX441176.1 GI:21665758
KEYWORDS
SOURCE synthetic construct
ORGANISM synthetic construct
other sequences; artificial sequences.
REFERENCE
1 Sallberg, M. and Hultgren, C.
AUTHORS Vaccines containing ribavirin and methods of use thereof
TITLE Patent: WO 0213855-A 16 21-FEB-2002;
JOURNAL TRIPEP AB (SE)
FEATURES
location/Qualifiers
source 1..2061
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/note="Hepatitis C virus NS3/4A coding region"
ORIGIN
Alignment Scores:
Pred. No.: 7,166-203 Length: 2061
Score: 3618.00 Matches: 686
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0
US-09-930-591-2 (1-686) x AX441176 (1-2061)
Qy 1 MetAlaProIleThrAlaTyAlaGlnGlnThrArgGlyLeuGlyCysIleIleThr 20
Db 1 ATGGCGCCTATCAGCGCCTATGCCAGCAGACAGAGGGCCCTTTGGGATGCATATCACC 60
Qy 21 SerLeuThrGlyArgAspLysAsnGlnValGluGlyGluValGlnIleValSerThra 40
Db 61 AGCTTGACCGCGCGGACAAACACAGGTGAGGGTGAGGTTTCAGATCGTGTCAACTGCT 120
Qy 41 AlaGlnThrPheLeuAlaThrCysIleAsnGlyValCysTrpThrValTyHisGlyAla 60
Db 121 GCCCAGACTTCTTGGCAACCTGCATTAAACGGGGTGTGTGACTGTCTACCATGGAGCC 180
Qy 61 GlyThrArgThrIleAlaSerProLysGlyProValIleGlnMetTyrThrAsnValAsp 80
Db 181 GGACACGAGACCATGTCCTTACCTAAGGTCCTGTATTCCAGATGTACCAATGTGGAC 240
Qy 81 GlnAspLeuValGlyTyrProAlaProGlnGlnGlyAlaArgSerLeuThrProCysThrCys 100
Db 241 CAAGACCTCGTAGCTGGCGCGCTCCCAAGGTGCGCGCTCATTAACACCATGCACATTGC 300
Qy 101 GlySerSerAspLeuTyrLeuValThrArgHisAlaAspValIleProValArgArgArg 120
Db 301 GGCTCCTCGGACCTTTACTGTCGTCACGAGGACGCGCATGTCAATTCCTGTGCGCGCAGG 360
Qy 121 GlyAspGlyArgGlySerLeuLeuSerProArgProIleSerTyrLeuLysGlySerSer 140
Db 361 GGTGATGCGAGGGCAGCGCTGCTTTCGCCCGCGCTATCTCTTAAAGGCTCCTCG 420
Qy 141 GlyGlyProLeuLeuCysProAlaGlyHisAlaValGlyIlePheArgAlaAlaValCys 160
Db 421 GGAGGCCCTCTGCTGTGCGCGCAGGACATGCCGTAGGCATATTACAGAGCGCGGTATGC 480
Qy 161 ThrArgGlyValAlaLysAlaValAspPheIleProValGluSerLeuGluThrThrMet: 180
Db 481 ACCCGTGAGTGGGTGAAGCGGTGACTTCATCCCGTAGAGCTTACAGACCAACCATG 540
Qy 181 ArgSerProValPheSerAspAsnSerSerProProAlaValProGlnSerTyrGlnVal 200
Db 541 AGGTCCCGGTGTCTCAGAACCTCTCCACACAGCAGTGCACAGTGCACCAAGTG 600
201 AlaHisIleuHisAlaProThrGlySerGlyLysSerThrLysValProAlaAlaTyrAla 220
601 GCCCACCTGCTATGCTCCACCGGAGGGTAAGAGCACCAAGGTCCCGCGGCATACGCA 660
221 AlaGlnGlyTyrLysValLeuValLeuAsnProSerValAlaAlaThrMetGlyPheGly 240
661 GCTCAGGGCTACAAGGTGCTGCTCAACCCCTCGTGTCTGCAACAAATGGGCTTTGGT 720
241 AlaTyrMetSerLysAlaHisGlyIleAspProAsnIleArgThrGlyValArgThrIle 260
721 GCTTACATGTCCAAGGCCCATGGATTGATCTTAACATCAGACTGGGGTCAGGACAATT 780
261 ThrThrGlySerProIleThrTyrSerThrTyrGlyLysPheLeuAlaAspGlyCys 280
781 ACTACTGGCAGCCCGATCAGCTATTCCACCTACGCGAAGTTCTTGGCAGCGGGGTGT 840
281 SerGlyGlyAlaTyrAspIleIleCysAspGluCysHisSerThrAspAlaThrSer 300
841 TCAGGGGTGCTTATGACATAAATAATTTGTACAGGTGCCACTCCACGATGCAACATCC 900
301 IleLeuGlyIleGlyThrValLeuAspGlnAlaGluThrAlaGlyAlaArgLeuThrVal 320
901 ATCTTGGGCATTGGCACTGCTTTCACAGCAGACCGCGGGCGGAGACTGACTGTG 960
321 LeuAlaThrAlaThrProGlySerValThrValProHisProAsnIleGluGluVal 340
961 CTCGCCACCGCTACCCCTCGGGCTCGGTCACTGTGCCCATCTCTAAATCGAGAGGTT 1020
341 AlaLeuSerThrThrGlyGluIleProPheTyrGlyLysAlaIleProLeuGluAlaIle 360
1021 GCTCTGTCCACTACCGGAGAGATCCCTTTTATGGCAAGGCTATTTCCTCTTGAAGCAAT 1080
361 LysGlyGlyArgHisLeuIlePheCysHisSerLysLysCysAspGluLeuAlaAla 380
1081 AAGGGGGGGAGACATCTCATCTTCTGCCACTCAAGAAGAGTGGCAGACTGCGCGCA 1140
381 LysLeuValAlaLeuGlyValAsnAlaValAlaTyrTyrArgGlyLeuAspValSerVal 400
1141 AAATCGTTCGGTGGGCGTCAATCGCTGCTTACTACCGCGGCTTGTGTGTGTCGTC 1200
401 IleProThrSerGlyAspValValAlaThrAspAlaLeuMetThrGlyPheThr 420
1201 ATCCCGACCATGTTGTCAGTTCGTGTCGCAACTGACGCCCTCATGACCGGCTTTACC 1260
421 GlyAspPheAspSerValIleAspCysAsnThrCysValThrGlnThrValAspPheSer 440
1261 GCGCACTTCGATTCGGTGTAGACTGCACACGCTGTGTCCACGAGACAGTGCATTCAGC 1320
441 LeuAspProThrPheThrIleGluThrIleThrLeuProGlnAspAlaValSerArgThr 460
1321 CTTGACCCCTACCTTCACCATTTAGACAATCAGCTTCCCAAGGATGCTGTCTCCGCTACT 1380
461 GlnArgArgGlyArgThrGlyArgGlyAspProGlyIleTyrArgPheValAlaProGly 480
1381 CAACGTGCGGGTAGGACTGGCAGAGGAGCAGGCATCTACAGATTGTGTGCGACCGGGG 1440
481 GluArgProSerGlyMetPheAspSerSerValLeuCysGluCysTyrAspAlaGlyCys 500
1441 GAGCGTCTCTCTGCGATGTTGACTCGTCTGCTCTCTGCGAGTGTATGACGCGGGTGT 1500
501 AlaTrpTyrGluLeuThrProAlaGluThrThrValArgLeuArgAlaTyrMetAsnThr 520
1501 GCTTGGTATGAGCTTACGCCCGCGGAGACACACAGCTTAGGCTACGAGCATACATGAACACC 1560
521 ProGlyLeuProValCysGlnAspHisIleuGluPheTrpGluGlyValPheThrGlyLeu 540
1561 CCGGGACTTCGCTGTGCCAAGACCATCTTCAATTTTGGGAGGCGCTTTTACGGGTCTC 1620
541 ThrHisIleAspAlaHisPheLeuSerGlnThrLysGlnSerGlyGluAsnLeuProTyr 560
1621 ACCCAGATAGACGCCCATCTCTTATCCAGACAAAGCAGAGTGGGGAAAACTTCCCTTAT 1680
561 LeuValAlaTyrGlnAlaThrValCysAlaArgAlaGlnAlaProProSerTrpAsp 580
```

```
Db      1681  CTGGTAGGTACCAAGCCCGGTGGCTAGAGCTCAAGCCCTCCCGCGTGGGAC 1740
Qy      581  GlnMetTrpLysCysLeuIleArgLeuLysProThrLeuHisGlyProThrProLeuLeu 600
Db      1741  CAGATGTGGAAGTGTGTATCCGTCTCAAGCCACCCCTCCATGGGCAACACCTCTGCTA 1800
Qy      601  TyrArgLeuGlyAlaValGlnAsnGluValThrLeuThrHisProValThrLysTyrIle 620
Db      1801  TATAGACTGGGCGCTGTCCAGATGAAGTCAACCTCAGCGCACCGATGACCAAGTATATC 1860
Qy      621  MetThrCysMetSerAlaAspLeuGluValValThrSerThrTrpValLeuValGlyGly 640
Db      1861  ATGACATGTATGTCCGCTGACCTGGAGGTGCTCAGGATACCTGGGTGCTGGTGGCGGC 1920
Qy      641  ValLeuAlaAlaLeuAlaLafyrCysLeuSerThrGlyCysValValIleValGlyArg 660
Db      1921  GTTCTGGCTCTTTGGCGCGTATTGCTATTCACAGGCTGCTGCTCATAGTAGTAGG 1980
Qy      661  IleValLeuSerGlyLysProAlaIleIleProAspArgGluValLeuTyrArgGluPhe 680
Db      1981  ATTGTCTGTTCGGAAGCGGCAATCATACCCGACAGGGAAGTCTCTACCGGGAGTTC 2040
Qy      681  AspGluMetGluGluCys 686
Db      2041  GATCAATGGAAGGTGC 2058

RESULT 3
AX467113          2061 bp      DNA      linear      PAT 17-JUL-2002
LOCUS             AX467113
DEFINITION        Sequence 1 from Patent WO0214362.
ACCESSION         AX467113
VERSION           AX467113.1  GI:21900425
KEYWORDS           synthetic construct
SOURCE             synthetic construct
ORGANISM           other sequences; artificial sequences.
REFERENCE          1
AUTHORS            Sallberg,M.
TITLE              A hepatitis c virus non-structural ns3/4a fusion gene
JOURNAL            Patent: WO 0214362-A 1 21-FEB-2002;
                   TRIPEP AB (SE)
FEATURES           Location/Qualifiers
                   1..2061
                   /organism="synthetic construct"
                   /mol_type="unassigned DNA"
                   /db_xref="taxon:32630"
                   /note="Hepatitis C virus NS3/4A coding region"
ORIGIN
Alignment Scores:
Pred. No.:        7,15e-203      Length:      2061
Score:            3618.00        Matches:      686
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%  Mismatches:    0
Query Match:      100.00%      Indels:       0
DB:               6              Gaps:        0

US-09-930-591-2 (1-686) x AX467113 (1-2061)
Qy      1  MetAlaProIleThrAlaTyrAlaGlnGlnThrArgGlyLeuLeuGlyCysIleIleThr 20
Db      1  ATGGGCGCTATACGGCCCTATGCCAGCAGACAAGGGCGCTTTTGGGATGCATAATCAC 60
Qy      21  SerLeuThrGlyArgAspLysAsnGlnValGluGlyValGlnIleValSerThrAla 40
Db      61  AGCTTGACCGCGCGGACAAAACAGGTGGAGGGTGAGGTTGAGATCGTGTCAACTGCT 120
Qy      41  AlaGlnThrPheLeuAlaThrCysIleAsnGlyValCysTyrThrValTyrHisGlyVal 60
Db      121  GCCCAGACTTCTTGGGCACTTACCTGATTAACGGGGTGTGTGATGCTACCATGGAGCC 180
Qy      61  GlyThrArgThrIleAlaSerProLysGlyProValIleGlnMetTyrThrAsnValAsp 80
```

```
Db      181  GGAACAAGGACCATTTGCGTCACTAAGGTCCTGTATTCCAGATGTACCAATGTGGAC 240
Qy      81  GlnAspLeuValGlyTrpProAlaProGlnGlyAlaArgSerLeuThrProCysThrCys 100
Db      241  CAAGACCTCGTAGGCTGGCCCGCTCCCAAGGTGCCCGCTCATTTAACCACTGACATGCG 300
Qy      101  GlySerSerAspLeuTyrLeuValThrArgHisAlaAspValIleProValArgArg 120
Db      301  GGCTCTCTGGACCTTTACTGTGTACAGAGGACGCGGATGTCTATTCTTGAAGAGCTCCTCG 360
Qy      121  GlyAspGlyArgGlySerLeuLeuSerProArgProIleSerTyrIleuLysGlySerSer 140
Db      361  GGTGATGGCAGGGGACGCTGCTTTTCGCCCGCGCTATCTCTTACTTGAAGAGCTCCTCG 420
Qy      141  GlyGlyProLeuLeuCysProAlaGlyHisAlaValGlyIlePheArgAlaAlaValCys 160
Db      421  GGAGGCGCTCTGCTGTGCCCCCGCAGACATGCGGTAGGCAATTTACAGCCCGGTATGCG 480
Qy      161  ThrArgGlyValAlaLysAlaValAspPheIleProValGluSerLeuGluThrThrMet 180
Db      481  ACCCGTGGAGTGGCTAAGGCGGTGGACTTTCATCCCGGTAGAGAGCTTAGAGACAACATG 540
Qy      181  ArgSerProValPheSerAspAsnSerSerProProAlaValProGlnSerTyrGlnVal 200
Db      541  AGGTCCCGCGTGTCTCTCAGACAACCTCTCCCGCAGCAGTCCCGCAGAGCTTACCAAGTG 600
Qy      201  AlaHisLeuHisAlaProThrGlySerGlyLysSerThrLysValProAlaAlaTyrAla 220
Db      601  GCCCACTCGATGCTCCACCGGCGGTAGAGCACCAAGGTCCCGCGCGCATACGCA 660
Qy      221  AlaGlnGlyTyrLysValLeuValLeuAsnProSerValAlaAlaThrMetGlyPheGly 240
Db      661  GCTCAGGGCTACAAGGTGCTGTGTGCTCAACCCCTCGTTCCTGCTCAACAATTTGGCT 720
Qy      241  AlaTyrMetSerLysAlaHisGlyIleAspProAsnIleArgThrGlyValArgThrIle 260
Db      721  GCTTACATGTCCAAGGCCCATGGATGTGATTCCTTAACATCAGGACTGGGTGAGACAAT 780
Qy      261  ThrThrGlySerProIleThrTyrSerThrTyrGlyLysPheLeuAlaAspGlyGlyCys 280
Db      781  ACTACTGGCAGCCCGATCAGTATTCCACTACGGCAAGTTCCTTGCCGACGCGGGTGT 840
Qy      281  SerGlyGlyAlaTyrAspIleIleCysAspGluCysHisSerThrAspAlaThrSer 300
Db      841  TCAGGGGTGCTTATGACATAATAATTGTGACAGTGCCTCCACGAGTCAACATCC 900
Qy      301  IleLeuGlyIleGlyThrValLeuAspGlnAlaGluThrAlaGlyAlaArgLeuThrVal 320
Db      901  ATCTTGGGCATTGGCCTGTCTTGACCAAGCAGACCGCGGGGCGAGACTGACTGTG 960
Qy      321  LeuAlaThrAlaThrProGlySerValThrValProHisProAsnIleGluGluVal 340
Db      961  CTGCCACCGCTACCCCTCCGGGCTCGGTCACTGTGCCCCCATCTTAACATCGAGGAGTT 1020
Qy      341  AlaLeuSerThrThrGlyGluIleProPheTyrGlyLysAlaIleProLeuGluAlaIle 360
Db      1021  GCTCTGTCCACTACCGAGAGATCCCTTTTATGGCAAGGCTATTCCTCTTGAAGCAAT 1080
Qy      361  LysGlyGlyArgHisIleuIlePheCysHisSerLysLysCysAspGluLeuAlaAla 380
Db      1081  AAGGGGGGAGAGCATCTCATCTTCTGCCACTCAAGAAGAGTGCAGCAGCTCGCGCA 1140
Qy      381  LysLeuValAlaLeuGlyValAsnAlaValAlaTyrTyrArgGlyLeuAspValSerVal 400
Db      1141  AACTGGTCCGCTGGGCGTCAATCGGTGCTTACTACCGCGCTTGTATGTGTCGTC 1200
Qy      401  IleProThrSerGlyAspValValValAlaThrAspAlaLeuMetThrGlyPheThr 420
Db      1201  ATCCCGACCACTGTGTCGCTGTGCGCAACTGACGCCCTCATGACCGGCTTTACC 1260
Qy      421  GlyAspPheAspSerValIleAspCysAsnThrCysValThrGlnThrValAspPheSer 440
```

```
Db 1261 GCGGACTTCGATTCGGTATAGACTGCAACACGCTGTGTACCCAGACAGTCGACTTCAGC 1320
Qy 441 LeuaspProThrPheThrIleGluThrIleThrLeuProGlnAspAlaValSerArgThr 460
Db 1321 CTTGACCCCTACCTTCACCATGAGACAATACGCTTCCCGAGGATGCTGTCTCCCGTACT 1380
Qy 461 GlnArgArgGlyArgThrGlyArgGlySerProGlyIleTyrArgPheValAlaProGly 480
Db 1391 CAAGCTCGGGGTAGAGCTGGCAGAGGGAAGCCAGCATCTACAGATTGTGGCACCCGGGG 1440
Qy 481 GluArgProSerGlyMetPheAspSerSerValLeuLeuGlyCysGlyTyrAspAlaGlyCys 500
Db 1441 GAGCGTCTCTTGGCACTGTTGACTCGTCTGCTCTCTCGAGTGTATGAGCGCGGTGT 1500
Qy 501 AlaTrpTyrGluLeuThrProAlaGluThrThrValArgLeuArgAlaTyrMetAsnThr 520
Db 1501 GCTTGGTATGAGCTTACCGCCGCCGAGACACAGTAGCTAGCTACGAGCATACATGAACACC 1560
Qy 521 ProGlyLeuProValCysGlnAspHisLeuGluPheTrpGluGlyValPheThrGlyLeu 540
Db 1561 CCGGAGCTTCCCGTGTGCCAGACCACTTGAATTTTGGGAGGGCGTCTTTACCGGGTCTC 1620
Qy 541 ThrHisIleAspAlaHisPheLeuSerGlnThrLysGlnSerGlyGluAsnLeuProTyr 560
Db 1621 ACCACATAGAGCGCCACTTCTATCCAGCAAGACGAGTGGGGAACCTTCCCTAT 1680
Qy 561 LeuValAlaTyrGlnAlaThrValCysAlaArgAlaGlnAlaProProSerTrpAsp 580
Db 1681 CTGCTAGCGTACCAAGCCACCGTGTGGCTAGAGCTCAAGCCCTCCCGCTGTGGGAC 1740
Qy 581 GlnMetTrpLysCysLeuIleArgLeuLysProThrLeuHisGlyProThrProLeuLeu 600
Db 1741 CAGATGTGGAAGTGTGTCATCGCTCAAGCCACCTCCATGGGCCAACACCTCTGCTA 1800
Qy 601 TyrArgLeuGlyAlaValGlnAsnGluValThrLeuThrHisProValThrLysTyrIle 620
Db 1801 TATAGATGGCGCGCTGTCCAGAAAGTCACTACCTGACGCCACCCAGTCACCAAGTATATC 1860
Qy 621 MetThrCysMetSerAlaAspLeuGluValValThrSerThrTrpValLeuValGlyGly 640
Db 1861 ATGACATGATGTTCGGCTGACCTGGAGTCTCAGAGTACTCTGGTGTCTGTGGCGGC 1920
Qy 641 ValLeuAlaLeuAlaLeuAlaTyrCysLeuSerThrGlyCysValValIleValGlyArg 660
Db 1921 GTTCTGGCTGTTTGGCGCGCTATTGCTATCCACAGGCTCGTGGTGCATAGTAGTAGG 1980
Qy 661 IleValLeuSerGlyLysProAlaIleIleProaspArgGluValLeuTyrArgGluPhe 680
Db 1981 ATTGTCTGTCCGAAAGCCGGCAATCATACCCGACAGGGAAGTCTCTACCGGGAGTTC 2040
Qy 681 AspGluMetGluGluCys 686
Db 2041 GATGAATGGAGAGTGC 2058
```

```
RESULT 4
CQ827032
LOCUS CQ827032 2078 bp DNA linear PAT 29-JUN-2004
DEFINITION Sequence 35 from Patent WO2004048402.
ACCESSION CQ827032
VERSION CQ827032.1 GI:49455663
KEYWORDS
SOURCE
ORGANISM synthetic construct
other sequences; artificial sequences.
REFERENCE 1
AUTHORS Sallberg,M.
TITLE A hepatitis c virus codon optimized non-structural ns3/4a fusion
gene
JOURNAL Patent: WO 2004048402-A 35 10-JUN-2004;
AUTHORS TRIPEP AB (SE)
FEATURES
location/Qualifiers
1..2078
/organism="synthetic construct"
```

```
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/notes="Codon optimized hepatitis C virus NS3/4A coding region"
12..2072
/notes="unnamed protein product"
/codon_start=1
/transl_table=11
/protein_id="CAG44601.1"
/db_xref="GI:49455664"
/transl_table=11
/translation="MAPITAAQOTRGLLCIITSLTRDKNOVEGVOIVSTAOTF
LATCINGCVTVVHGAGCTRTIASPKGPVIOMYTNVDOLVGMPPAQGARSITPTCCGS
SDLYLVRHADVIVRRRGDRGSLSPRISYLYKSGSGGLPCPHAGHVGIFRAAVC
TRGAKAVDFIPVESLETTMRSPVSDNSPPAVPQSYQVAHLHAPTSKSTKVPAA
YAAQGYKVLNFSVAATMGFAYMSKAHGDIPNRTGARTITTSITTSYTGKFLA
DNTEVALSTTGTIPFYGKAIPLAIEKGRHLIFCHSKKCDLAALVALGNVAV
PGLDVSIVPTSGDVVVAVDALMTGFTGDPDSVIDCNTCVTQTVDSLDPTFLETI
TLQDASVTRQRRGRTGRKGIYRVPAPERPSGMFSSVLCSCDAGCAWYELTIPA
ETTVRLRYNNTTGLPVCQDHLFEWGFVFTGLTHIDAHFLSQTKOSGENLPYLVAQA
TVCAQAQAPPPSDMMWKILRLKPTLHGFTPLLYRLGAVQNEVTLTHPTVTKIMTCM
SADLEVVTWVLVGGVLAALAAAYCLSTGCVIVGRIVLVSGLPAIIPDREVLREFDE
MEEC"
```

## ORIGIN

```
Alignment Scores:
Pred. No.: 7,23e-203 Length: 2078
Score: 3618.00 Matches: 686
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0
```

US-09-930-591-2 (1-686) x CQ827032 (1-2078)

```
Qy 1 MetAlaProIleThrAlaTyrAlaGlnGlnThrArgGlyLeuLeuGlyCysIleIleThr 20
Db 12 ATGCCCCCATACCGCTACCGCCAGACACCGCGCGCTGCTGGGTGCATCATCACC 71
Qy 21 SerLeuThrGlyArgAspLysAsnGlnValGluGlyGluValGlnIleValSerThrAla 40
Db 72 AGCTGACCGCGCCGCAAGAACACAGGTGGAGGGCGAGGTGCAGATCGTGAGCACCGCC 131
Qy 41 AlaGlnThrPheLeuAlaThrCysIleAsnGlyValCysTrpThrValTyrHisGlyAla 60
Db 132 GCCAGACCTTCTCGGCCACCTGCATCAACGGCGTGTGCTGACCGGTACACCGCGCC 191
Qy 61 GlyThrArgThrIleAlaSerProLysGlyProValIleGlnMetTyrThrAsnValAsp 80
Db 192 GGCACCGCACCATCGCCAGCCCGCCCGGCGCGCTGATCCAGATGTACACCAACGTTGAC 251
Qy 81 GlnAspLeuValGlyTrpProAlaProGlnGlyAlaArgSerLeuThrProCysThrCys 100
Db 252 CAGGACCTGTGTGGCTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 311
Qy 101 GlySerSerAspLeuTyrLeuValThrArgHisAlaAspValIleProValArgArgArg 120
Db 312 GGCAGCAGCGACCTGTACTTGGTGACCGCGCACCGCGCGCGCGCGCGCGCGCGCGCG 371
Qy 121 GlyAspGlyArgGlySerLeuLeuSerProArgProIleSerTyrLeuLysGlySerSer 140
Db 372 GGCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 431
Qy 141 GlyGlyProLeuLeuCysProAlaGlyHisAlaValGlyIlePheArgAlaAlaValCys 160
Db 432 GGCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 491
Qy 161 ThrArgGlyValAlaLysAlaValAspPheIleProValGluSerLeuGluThrMet 180
Db 492 ACCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 551
Qy 181 ArgSerProValPheSerAspAsnSerSerProProAlaValProGlnSerTyrGlnVal 200
```

```
Db 552 CGAGCGCCGCTGTTACGACGACAAACAGCAGCGCCCCCGCGCTGCCCGCAGAGCTACCGAGTG 611
Qy 201 AlaHisLeuHisAlaProThrGlySerGlyLysSerThrLysValProAlaAlaTyAla 220
Db 612 GCCACCTGACGCGCCCGCCAGCGGCGGACAGACACCAAGGTGCCCGCGCTACGCC 671
Qy 221 AlaGlnGlyTyrLysValLeuValLeuAsnProSerValAlaAlaThrMetGlyPheGly 240
Db 672 GCCAGGGCTACAAAGGTGCTGGTGTGAACCCAGCGGTGCCGCCACCATGGCTTCGGC 731
Qy 241 AlaTyrMetSerLysAlaHisGlyIleAspProAsnIleArgThrGlyValArgThrIle 260
Db 732 GCCTACATGACGAAGCGCCACGCGCATCGACCCCAACATCCGCGCGTGCACCATC 791
Qy 261 ThrThrGlySerProIleThrTyrSerThrTyrGlyLysPheLeuAlaAspGlyGlyCys 280
Db 792 ACCACGGGAGCCCATCATCTACAGCACTTACGGCAGTTCTTGGCCGACGCGGCTGC 851
Qy 281 SerGlyGlyAlaTyrAspIleIleCysAspGluCysHisSerThrAspAlaThrSer 300
Db 852 AGCGGGCGGCGCTACGACATCATCTCTCGACGAGTGCACAGCACCGACGCGCACCGAC 911
Qy 301 IleLeuGlyIleGlyThrValLeuAspGlnAlaGluThrAlaGlyAlaArgLeuThrVal 320
Db 912 ATCTTGGGCACTCGGCACCGCTGTGGACAGCGCGGAGACCGCGCGCGCCGCTGACCGTG 971
Qy 321 LeuAlaThrAlaThrProGlySerValThrValProHisProAsnIleGluGluVal 340
Db 972 CTGGCCACCGCCACCCCGCGGAGGTGACCGTGCCTGCCCGCCACCCCAACATCGAGGAGTG 1031
Qy 341 AlaLeuSerThrThrGlyGluIleProPheTyrGlyLysAlaIleProLeuGluAlaIle 360
Db 1032 GCCTGAGCACCAACCGCGGAGATCCCTTCTACGGCAAGGCCATCCCTCGGAGGCCATC 1091
Qy 361 LysGlyGlyArgHisLeuIlePheCysHisSerLysLysCysAspGluLeuAlaAla 380
Db 1092 AAGGGCGGCGCCACCTGTATCTTCTGCCACAGCAAGAAGTGCAGCAGCTGGCGGCC 1151
Qy 381 LysLeuValAlaLeuGlyValAsnAlaValAlaTyrTyrArgGlyLeuAspValSerVal 400
Db 1152 AAGCTGTGGCCCTTGGCGGTGAACCGCGTGCCTTACTACCGCGCGCTTGACGTGAGCGTG 1211
Qy 401 IleProThrSerGlyAspValValValAlaThrAspAlaLeuMetThrGlyPheThr 420
Db 1212 ATCCCAACACGCGGACGCGTGTGTGTGCCACCGACCGCTGATGACCGGCTTCAAC 1271
Qy 421 GlyAspPheAspSerValIleAspCysAsnThrCysValThrGlnThrValAspPheSer 440
Db 1272 GCGGACTTCGACAGCGGTGATCGACTGCAACACCTGCGTGACCCAGACCGGTGACCTT 1331
Qy 441 LeuAspProThrPheThrIleGluThrIleThrLeuProGlnAspAlaValSerArgThr 460
Db 1332 CTGGACCCCACTTACCATCGAGACCATCACCTGCGCCCGAGCGCGCTGACCGCGCAC 1391
Qy 461 GlnArgArgGlyArgThrGlyArgGlyPheProGlyIleTyrArgPheValAlaProGly 480
Db 1392 CAGCGCGCGCGCGCACCGCGCGCGGCAAGCCCGCATCTACCGCTTCTGCGCGCCCGCG 1451
Qy 481 GluArgProSerGlyMetPheAspSerSerValLeuCysGluCysTyrAspAlaGlyCys 500
Db 1452 GAGCGCGCCACGCGCATGTTTGAACAGCAGCGGTGCTGTGCAGTGTGTACGACGCGCGTGC 1511
Qy 501 AlaTyrTyrGluLeuThrProAlaGluThrThrValArgLeuArgAlaTyrMetAsnThr 520
Db 1512 GCCTGGTACGAGCTGACCCCGCGGAGACCATCCGCTGGCGCTGCGCGCTTACATGAACAC 1571
Qy 521 ProGlyLeuProValCysGlnAspHisLeuGluPheThrGluGlyValPheThrGlyLeu 540
Db 1572 CCGCGGCTGCGCGGTGCGCAGGACCATCTGAGTTCTTGGGAGGCGGTTCACCGGCGCTG 1631
Qy 541 ThrHisIleAspAlaHisPheLeuSerGlnThrLysGlnSerGlyGluAsnLeuProTyr 560
Db 1632 ACCACATCGACGCCCACTTCTGAGCCAGACCAAGCAGCGCGCGGACCTTGGCCCTAC 1691
```

```
Qy 561 LeuValAlaTyrGlnAlaThrValCysAlaArgAlaGlnAlaProProProSerTyrAsp 580
Db 1692 CTGGTGGCTACACAGGCCACCGCTGTGCGCCCGCGCCAGGCCCGCCCGCCAGCTGGAC 1751
Qy 581 GlnMetTyrLysCysLeuIleArgLeuLysProThrLeuHisGlyProThrProLeuLeu 600
Db 1752 CAGATGTGAAGTGCCTGATCCGCTTGAAGCCACCTGACGCGGCCACCCCGCTGCTG 1811
Qy 601 TyrArgLeuGlyAlaValGlnAsnGluValThrLeuThrHisProValThrLysTyrIle 620
Db 1812 TACCGCTGGCGCGCTGTGAGAACGAGGTGACCTGACCCCGCTGACCAAGTACATC 1871
Qy 621 MetThrCysMetSerAlaAspLeuGluValValThrSerThrTyrValLeuValGly 640
Db 1872 ATGACCTGCATGAGCGCGCCGCTGAGGTGTGTGACCAACCTGCTGTGGTGGCGGC 1931
Qy 641 ValLeuAlaAlaLeuAlaTyrCysLeuSerThrGlyCysValValIleValGlyArg 660
Db 1932 GTGCTGGCGCGCTGCGCTACTGCTGAGCACCGCTGCTGTGTGTGCTGTGGCGCG 1991
Qy 661 IleValLeuSerGlyLysProAlaIleIleProAspArgGluValLeuTyrArgGluPhe 680
Db 1992 ATGCTGCTGAGCGCGCAAGCCCGCATCTCCCGACCGAGGTGCTGTACCGGAGTTC 2051
Qy 681 AspGluMetGluGluCys 686
Db 2052 GACGAGATGGAGGAGTGTC 2069

RESULT 5
LOCUS AY615798
DEFINITION Hepatitis C virus isolate HCV.TWB.1 polyprotein gene, partial cds.
ACCESSION AY615798
VERSION AY615798.1 GI:48479029
KEYWORDS Hepatitis C virus
ORGANISM Hepatitis C virus
VIRUSES: sRNA positive-strand viruses, no DNA stage; Flaviviridae;
Hepacivirus.
REFERENCE 1 (bases 1 to 8791)
AUTHORS Brann,T.W., Kottlilil,S., Polis,M. and Imamichi,T.
TITLE Identification of mutations associated with interferon resistance in HCV and HIV co-infected patients
JOURNAL Unpublished
AUTHORS 2 (bases 1 to 8791)
Brann,T.W., Kottlilil,S., Polis,M. and Imamichi,T.
TITLE Direct Submission
JOURNAL Submitted (03-MAY-2004) LHR/CSP, SAIC-Frederick, Inc, Building 550, Room 126, 1050 Boyles Street, Frederick, MD 21702, USA
FEATURES
source
1..8791
/organism="Hepatitis C virus"
/mol_type="genomic RNA"
/strain="HCV1a"
/isolate="HCV.TWB.1"
/db_xref="taxon:11103"
/country="USA"
67..>8791
/codon_start=1
/product="polyprotein"
/protein_id="AAT44836.1"
/db_xref="GI:48479030"
/tranlation="MSTNPKPQKTKRNTNRPPQDVKFPGGQIVGGVLLPRRGRPL
GVRAIKTSRQPRGRROPFIKARPEGRTHAQPCYPMPLVGNCGGAGWLLSRG
SPSGMPTDPRRRNRLKVIDTLTCGFADLMGIPIPLVGAPLGGAARALAHGVRVLED
GVNYATGNCIPGCSFSLFLLALSCLTPASAYQVRNSGLYHVINDCPNSSIVYETAD
AILHSPGCIPCVREGNASKCWAVATPVRIDRGLRSTQLRRHIDLLVSSATLCSALY
VBDLCSVLVQLFTFSPRRHWTODNCISYPGHTGHRWAMDMMWNSPTTLVW
AQLRIPOAILDMIAGHWGLAGIAYFSMVGNWAKVLVLLFACVDATSTTGAAG
ATATVPTFFAPAGPKONVLIYTNGSWHINRTALNCNASLDTGWLALFYVNSFNSS
GCPERNASCRPLASFDQGWGPISYANGSPRHPYPCWHYPPKPCGLVPAQSGVCGPYC
FTFSPVVGTTDKFGVPTYNWANDTDVFLNNTNRPLGNWFGCTWMNSTGTFKVCGA
```







```
Db 4522 CAACGCGGGTAGGACCGGAGGGGAGCCAGGCATCTACAGATTGTGGCCCGGG 4581
Qy 481 GluArgProSerGlyMetPheAspSerSerValLeuCysGluCysTyrAspAlaGlyCys 500
Db 4582 GAGCGCCCTTCGCGCATGTTTGATCGTCCGTCCTCTGCGAGTCTATGACGCGGGTGT 4641
Qy 501 AlaTyrTyrGluLeuThrProAlaGluThrThrValArgLeuArgAlaTyrMetAsnThr 520
Db 4642 GCTTGGTATGAGCTCACACCGCCGAGACACAGTATAGGCTACAGCATCATGAACACC 4701
Qy 521 ProGlyLeuProValCysGlnAspHisLeuGluPheTyrGluGlyValPheThrGlyLeu 540
Db 4702 CCGGGGCTTCGCGTGTGCCAAGACCATCTTGAATTTTGGGAGGCGTTCACGGGTCT 4761
Qy 541 ThrHisLeuAspAlaHisPheLeuSerGlnThrIysGlnSerGlyGluAsnLeuProTyr 560
Db 4762 ACCCACATAGACGCCCACTTCCTATCCAGAACAGAGTGGGGAGAACCTTCCTTAC 4821
Qy 561 LeuValAlaTyrGlnAlaThrValCysAlaArgAlaGlnAlaProProSerTyrAsp 580
Db 4822 CTGGTGGCTACCAAGCACCGTGTGCTAGAGCTCAAGCCCCCCCCCGTGTGGGAC 4881
Qy 581 GlnMetTrpCysCysLeuIleArgLeuLysProThrLeuHisGlyProThrProLeuLeu 600
Db 4882 CAGATGTGAAGTCTTGATCCGCTCAAGCCACACTCCATGGGCCAACACCTCTGCTA 4941
Qy 601 TyrArgLeuGlyAlaValGlnAsnGluValThrLeuThrHisProValThrIysTyrIle 620
Db 4942 TACAGGCTGGGTGCGGTTCCAGATGAAGTCAACCTGACGACCACCAAGTCAACAGTACATC 5001
Qy 621 MetThrCysMetSerAlaAspLeuGluValValThrSerThrTrpValLeuValGlyGly 640
Db 5002 ATGACATGCATGTCGCGCGACCTGGAGTCTGTCAGAGTACCTGGGTGCTCGTGGTGGC 5061
Qy 641 ValLeuAlaLeuAlaTyrCysLeuSerThrGlyCysValValIleValGlyArg 660
Db 5062 GTTCTGGCTCTTTGGCGCGTACTGCCTATCCACAGGCTGCGTGTGTCATATAGGCAGG 5121
Qy 661 IleValLeuSerGlyLysProAlaIleProAspArgGluValLeuTyrArgGluPhe 680
Db 5122 ATTGTCTGTGCGGGAAGCGGAGTCATATCCCGCAGGAGGTCTCTACCGAGAGTTC 5181
Qy 681 AspGluMetGluGluCys 686
Db 5182 GATGAAATGGAAGAGTGC 5199

RESULT 6
LOCUS I06434 5360 bp DNA linear PAT 02-DEC-1994
DEFINITION Sequence 48 from Patent EP 0318216.
ACCESSION I06434
VERSION I06434.1 GI:590311
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE
1 (bases 1 to 5360)
AUTHORS Houghton,M., Choo,Q.-L. and Kuo,G.
TITLE Nanbv diagnostics and vaccines
JOURNAL Patent: EP 0318216-A1 48 31-MAY-1989;
FEATURES
Location/Qualifiers
1..5360
/organism="unknown"
/mol_type="unassigned DNA"
ORIGIN

Alignment Scores:
Pred. No.: 8.65e-200 Length: 5360
Score: 3574.00 Matches: 672
Percent Similarity: 99.42% Conservatative: 10
Best Local Similarity: 97.96% Mismatches: 4
Query Match: 98.78% Indels: 0
```

```
DB: 6 Caps: 0
US-09-930-591-2 (1-686) x I06434 (1-5360)
Qy 1 MetAlaProIleThrAlaTyrAlaGlnGlnThrArgGlyLeuLeuGlyCysIleIleThr 20
Db 930 CTGGCGCCCATCATCAGCGGTAAGCCAGCAGACAAAGGGGCTCTCTAGGTGCATAATCACC 989
Qy 21 SerLeuThrGlyArgAspLysAsnGlnValGluGlyGluValGlnIleValSerThrAla 40
Db 990 AGCTTAACCTGGCGGACAAAACCAAGTGGAGGTGAGGTCCAGATTGTGCAACTGCT 1049
Qy 41 AlaGlnThrPheLeuAlaThrCysIleAsnGlyValCysTrpThrValTyrHisGlyAla 60
Db 1050 GCCCAAACTTCTTGGCAACGTGATCAATGGGGTGTCTGGACTGTCTACCAAGGGGCC 1109
Qy 61 GlyThrArgThrIleAlaSerProLysGlyProValIleGlnMetTyrThrAsnValAsp 80
Db 1110 GGAACGAGGACCATCGCGTCACCAAGGTCCTGTCTCATCCAGATGTATACCAATGTAGAC 1169
Qy 81 GlnAspLeuValGlyTrpProAlaProGlnGlyAlaArgSerLeuThrProCysThrCys 100
Db 1170 CAAGACCTTGTGGGCTGGCCGCTCCGCAAGGTAGTGGCTCATTTGACACCTGCACTTGC 1229
Qy 101 GlySerSerAspLeuTyrLeuValThrArgHisAlaAspValIleProValArgArg 120
Db 1230 GGCTCTCCGACCTTTACCTGTGTACGAGGACGCGGATGTCTTCCCTGGCGCGCGG 1289
Qy 121 GlyAspGlyArgGlySerLeuLeuSerProArgProIleSerTyrIleuLysGlySerSer 140
Db 1290 GGTGTATAGCAGGGGAGCCTGTGTGCGCCCGGCCCATTTCTTCTTGAAGAGCTCTCG 1349
Qy 141 GlyGlyProLeuLeuCysProIleGlyHisAlaValGlyIlePheArgAlaAlaValCys 160
Db 1350 GGGGGTCCGCTGTGTGCCCCCGGGGACGCGGTGGGCATATTTAGGGCCGCGGTGTC 1409
Qy 161 ThrArgGlyValAlaLysAlaValAspPheIleProValGluSerLeuGluThrThrMet 180
Db 1410 ACCGTGGAGTGGCTAAGCGGTGGACTTTATCCCTGTGGAGAACTTAGAGCAACCATG 1469
Qy 181 ArgSerProValPheSerAspAsnSerSerProProAlaValProGlnSerTyrGlnVal 200
Db 1470 AGGTCCCGGCTTTCACGGATAACTCTCTCCACAGTAGTCCCGCAGAGCTTCCAGGTG 1529
Qy 201 AlaHisLeuHisAlaProThrGlySerGlyLysSerThrLysValProAlaAlaTyrAla 220
Db 1530 GCTCACTCCATGCTCCACAGCGAGGCAAAAGCACCAGGTCCCGCTGCATATGCA 1589
Qy 221 AlaGlnGlyTyrLysValLeuValLeuAsnProSerValAlaAlaThrMetGlyPheGly 240
Db 1590 GCTCAGGGCTATAAGGTGCTAGTACTCAACCCCTCTGTGTCGCAACACTGGGCTTGTG 1649
Qy 241 AlaTyrMetSerLysAlaHisGlyIleAspProAsnIleArgThrGlyValArgThrIle 260
Db 1650 GCTTACATGTCCAGGCTCATGGGATCGATCTTAACATCAGGACCGGGGTGAGAACAAAT 1709
Qy 261 ThrThrGlySerProIleThrTyrSerThrTyrGlyLysPheLeuAlaAspGlyCys 280
Db 1710 ACCACTGGCAGCCCCCATCAGCTACTCCACCTACGGCAAGTTCCTTGGCGACGGGGTGC 1769
Qy 281 SerGlyGlyAlaTyrAspIleIleCysAspGluCysHisSerThrAspAlaThrSer 300
Db 1770 TCGGGGGGCGCTTATGACATAATAATTTGTGACGAGTGCCACTCCACGGATGCCACATCC 1829
Qy 301 IleLeuGlyIleGlyThrValLeuAspGlnAlaGluThrAlaGlyAlaArgLeuThrVal 320
Db 1830 ATCTTGGGCATCGGCACCTGTCTTGAACAGCAGAGACTGGCGGGGCGAGACTGTTGTG 1889
Qy 321 LeuAlaThrAlaThrProGlySerValThrValProHisProAsnIleGluGluVal 340
Db 1890 CTCGCCACCGCCACCCCTCCGGGCTCCGCTCACTGTGCCCCCATCCCAACATCGAGAGTT 1949
Qy 341 AlaLeuSerThrThrGlyGluIleProPheTyrGlyLysAlaIleProLeuGluAlaIle 360
```

```
Db 1950 GCTCTGTCCACACCGGAGAGATCCCTTTTACGGCAAGGCTATCCCTCGAAGTAATC 2009
Qy 361 LysGlyGlyArgHisLeuIlePheCysHisSerLysLysCysAspGluLeuAlaAala 380
Db 2010 AAGGGGGGAGACATCTCTCTGTCTATCAAGAGAGAGAGTGCGAGCACTCGCGCA 2069
Qy 381 LysLeuValAlaLeuGlyValAsnAlaValAlaTyrTyrArgGlyLeuAspValSerVal 400
Db 2070 AAGTGTGTGGCATGGGATCAATCCGTCGCTACTACCGCGGTCTTGAGGTGTCCGTC 2129
Qy 401 IleProThrSerGlyAspValValValAlaThrAspAlaLeuMetThrGlyPheThr 420
Db 2130 ATCCGACCGGGGAGTGTGTGTGTGTGGCAACCGATGCCCTCATACCGGCTATACC 2189
Qy 421 GlyAspPheAspSerValIleAspCysAsnThrCysValThrGlnThrValAspPheSer 440
Db 2190 GCGGACTTCGACTCGGTGATAGACTGCAATACGTGTGTACCCGACAGTGCATTTACG 2249
Qy 441 LeuAspProThrPheThrIleGluThrIleThrLeuProGlnAspAlaValSerArgThr 460
Db 2250 CTTGACCCCTACCTTCACCATTTGAGACAATCAGCTCCCGCAGGATGCTGTCTCCGCACT 2309
Qy 461 GlnArgArgGlyArgThrGlyArgGlyLysProGlyIleTyrArgPheValAlaProGly 480
Db 2310 CAACGTCCGGGCGAGACTGGCAGGGGGAAGCCAGGCATCTACAGATTTGTGGCACCGGG 2369
Qy 481 GluArgProSerGlyMetPheAspSerSerValLeuCysGluCysTyrAspAlaGlyCys 500
Db 2370 GAGGCCCTCCGGCATGTTGCATCTGCTCCCTCTGTGAGTGTATGACGAGGCTGT 2429
Qy 501 AlaTrpTyrGluLeuThrProAlaGluThrThrValArgLeuArgAlaTyrMetAsnThr 520
Db 2430 GCTTGGTATGAGCTCACGCCCGCGAGACTACAGTTAGGCTACGAGGTACATGAACACC 2489
Qy 521 ProGlyLeuProValCysGlnAspHisLeuGluPheTrpGluGlyValPheThrGlyLeu 540
Db 2490 CCGGGGCTTCCTCGGTGTGCAGGACCATCTTGAATTTTGGAGGGCGCTTTTACAGGCTC 2549
Qy 541 ThrHisIleAspAlaHisPheLeuSerGlnThrLysGlnSerGlyGluAsnLeuProTyr 560
Db 2550 ACTCATATAGATGCCCATCTTCTATCCGACAAAGACAGTGGGAGAACCTTCTCTTAC 2609
Qy 561 LeuValAlaTyrGlnAlaThrValCysAlaArgAlaGlnAlaProProSerTrpAsp 580
Db 2610 CTGCTAGCGTACCAAGCACCGTGTGCGCTAGGCTCAGGCCCTCCGCCATCGTGGAC 2669
Qy 581 GlnMetTrpLysCysLeuIleArgLeuLysProThrLeuHisGlyProThrProLeuLeu 600
Db 2670 CAGATGTGGAGTGTGTTGATTGCGCTCAAGGCCACCCCTCCATGGGCCAACACCCCTGCTA 2729
Qy 601 TyrArgLeuGlyAlaValGlnAsnGluValThrLeuThrHisProValThrLysTyrIle 620
Db 2730 TACAGATGGCGCGCTGTTCAGAAATCAATACCTCAGCAGCACCTGGTGTCTGTTGGCGGC 2789
Qy 621 MetThrCysMetSerAlaAspLeuGluValValThrSerThrTrpValLeuValGlyGly 640
Db 2790 ATGACATGCATGTTCGGCGGACCTGGAGTGTCTACAGGACCTGGTGTCTGTTGGCGGC 2849
Qy 641 ValLeuAlaAlaLeuAlaAlaTyrCysLeuSerThrGlyCysValValIleValGlyArg 660
Db 2850 GTCTCGGCTGTTTGGCGCGGTATTGCTGTCAACAGGCTCGTGGTGCATAGTGGCGAGG 2909
Qy 661 IleValLeuSerGlyLysProAlaIleIleProAspArgGluValLeuTyrArgGluPhe 680
Db 2910 GTCTGCTGTTCGGGGAAGCGGCAATCATACCTCACAGGGAAGTCTCTACCGAGAGTTTC 2969
Qy 681 AspGluMetGluGluCys 686
Db 2970 GATGAGATGGNAGAGTGC 2987
```

RESULT 7  
I09328

```
LOCUS I09328 5360 bp DNA linear PAT 02-DEC-1994
DEFINITION Sequence 8 from Patent WO 8904669.
ACCESSION I09328
VERSION I09328.1 GI:587963
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE
1 (bases 1 to 5360)
AUTHORS Houghton, M., Choo, O.-K. and Kuo, G.
JOURNAL Patent: WO 8904669-A 8 01-JUN-1989;
FEATURES
Location/Qualifiers
1..5360
/organism="unknown"
/mol_type="unassigned DNA"

ORIGIN
Alignment Scores: 8.65e-200 Length: 5360
Pred. No.: 3574.00 Matches: 672
Score: 3574.00
Percent Similarity: 99.42% Conservative: 10
Best Local Similarity: 97.96% Mismatches: 4
Query Match: 98.78% Indels: 0
DB: 6 Gaps: 0

US-09-930-591-2 (1-686) x I09328 (1-5360)

Qy 1 MetAlaProIleThrAlaTyrAlaGlnThrArgGlyLeuGlyCysIleIleThr 20
Db 930 CTGGCGGCCCATCATCGGGGTACGCCCGACAGACAAAGGGGCTCTTAGGGTGATATATCACC 989
Qy 21 SerLeuThrGlyArgAspLysAsnGlnValGluGlyValCysTrpThrValThrAla 40
Db 990 AGCTTAATCGCGGGGACAAAACCAAGTGGAGGTCGAGGTCCAGATTGTGTCACTGCT 1049
Qy 41 AlaGlnThrPheLeuAlaThrCysIleAsnGlyValCysTrpThrValThrHisGlyAla 60
Db 1050 GCCCAACCTTCTCTGGCAACGTGCATCAATGGGTGTGTGACTGTCTACACGGGGCC 1109
Qy 61 GlyThrArgThrIleAlaSerProLysGlyProValIleGlnMetTyrThrAsnValAsp 80
Db 1110 GGAACGAGGACCATCGCGTCCACCAAGGGTCTCTCATCAGATGTATACCAATGTAGAC 1169
Qy 81 GlnAspLeuValGlyTrpProAlaProGlnGlyAlaArgSerLeuThrProCysThrCys 100
Db 1170 CAAGACCTTGTGGCTGGCGCTCCGCAAGGTAGCGCTCATTTGACACCTGACCTTGC 1229
Qy 101 GlySerSerAspLeuTyrLeuValThrArgHisAlaAspValIleProValArgArg 120
Db 1230 GGCTCTCGGACCTTTACCTGGTCCAGGACGCGCATGTCTATCCGTGCGCGCGCG 1289
Qy 121 GlyAspGlyArgGlySerLeuLeuSerProArgProIleSerTyrLeuLysGlySerSer 140
Db 1290 GGTGATAGAGGGGACGCTGTGTGCGCCCGGCCCATTTCTACTTTGAAGGCTCTCTCG 1349
Qy 141 GlyGlyProLeuLeuCysProAlaGlyHisAlaValGlyIlePheArgAlaAlaValCys 160
Db 1350 GGGGTTCCGTGTGTGTGTCGCCCGGGGACGCGCGTGGGCATATTTAGGCGCGCGGTGTC 1409
Qy 161 ThrArgGlyValAlaLysAlaValAspPheIleProValGluSerLeuGluThrThrMet 180
Db 1410 ACCGTGGAGTGGCTAAGCGGTGGACTTTTATCCCTGTGGAGAACCTTAGACACAACCATG 1469
Qy 181 ArgSerProValPheSerAspAsnSerSerProProAlaValProGlnSerTyrGlnVal 200
Db 1470 AGGTCCCGGTGTTCACGGATAACTCTCTCCACAGTAGTGTGCCCGCCAGAGTTCACAGTG 1529
Qy 201 AlaHisLeuHisAlaProThrGlySerGlyLysSerThrLysValProAlaAlaTyrAla 220
Db 1530 GCTCACCTCCATGCTCTCCACAGGCGGCAAGACCAAGGTCCTCCGCTGCATATGCA 1589
Qy 221 AlaGlnGlyTyrLysValLeuValLeuAsnProSerValAlaAlaThrMetGlyPheGly 240
```

Db 1590 GCTCAGGGCTATAGGTGCTAGTACTCAACCCCTCTGTTGCTGCAACACTGGGCTTTGGT 1649  
QY 241 AlaTyrMetSerLysAlaHisGlyIleAspProAsnIleArgThrGlyValArgThrIle 260  
Db 1650 GCTTACATGTCGAAGGCTCATGGATCGATCCTAACATCAGGACCGGGGTGAGAACAT 1709  
QY 261 ThrThrGlySerProIleThrTyrSerThrTyrGlyLysPheLeuAlaAspGlyGlyCys 280  
Db 1710 ACCACTGGCAGCCCACTACTCCACTACGGCAAGTTCTTCCCGACGCGGGTGC 1769  
QY 281 SerGlyGlyAlaTyrAspIleIleCysAspGluCysHisSerThrAspAlaThrSer 300  
Db 1770 TCGGGGGCGCTTATGACATAATTTGTGACGAGTGCCTCCACGATGCCATCC 1829  
QY 301 IleLeuGlyIleGlyThrValLeuAspGlnAlaGluThrAlaGlyAlaArgLeuThrVal 320  
Db 1830 ATCTTGGGCATCGGCATGTCCTTGACCAAGCAGAGACTCGGGGGGAGACTGGTTGTG 1889  
QY 321 LeuAlaThrAlaThrProProGlySerValThrValProHisProAsnIleGluGluVal 340  
Db 1890 CTCGCCACCGCCACCTCTCGGGCTCGTCACTGTGCCCATCCCAACATCGAGGAGTT 1949  
QY 341 AlaLeuSerThrThrGlyGluIleProPheTyrGlyLysAlaIleProLeuGluAlaIle 360  
Db 1950 GCTCTGTCCACCAACCGAGAGATCCCTTTTACGCAAGGCTATCCCTCTCGAAGTAATC 2009  
QY 361 LysGlyGlyArgHisLeuIlePheCysHisSerLysLysLysCysAspGluLeuAlaAla 380  
Db 2010 AAGGGGGGAGACATCTCATCTTCTGTCATTAAGAGAAAGTGCAGCACTCGCGCA 2069  
QY 381 LysLeuValAlaLeuGlyValAsnAlaValAlaTyrTyrArgGlyLeuAspValSerVal 400  
Db 2070 AACTGTGTGTCATTTGGCATCAATCGGTGGCTACTACCGCGTCTTGACGTGTCGTC 2129  
QY 401 IleProThrSerGlyAspValValValAlaThrAspAlaLeuMetThrGlyPheThr 420  
Db 2130 ATCCCGACCAACCGCGGATGTGTGCTGTGCGCAACCGATGCCCTCATGACCGGCTATAC 2189  
QY 421 GlyAspPheAspSerValIleAspCysAsnThrCysValThrGlnThrValAspPheSer 440  
Db 2190 GCGGACTTTCGACTCGGTGATAGACTGCATACGTGTGTACCCAGACAGTCGATTCAGC 2249  
QY 441 LeuAspProThrPheThrIleGluThrIleThrLeuProGlnAspAlaValSerArgThr 460  
Db 2250 CTTGACCTTACCTTACCATTAGACAAATCAGCTCCCCAGGATGTGTCTCCGCACT 2309  
QY 461 GlnArgArgGlyArgThrGlyArgGlyLysProGlyIleTyrArgPheValAlaProGly 480  
Db 2310 CAACGTCGGGCGAGACTGGCAGGGGAGCCAGGCATCTACAGATTTGTGGCACCAGGG 2369  
QY 481 GluArgProSerGlyMetPheAspSerSerValLeuCysGluCysTyrAspAlaGlyCys 500  
Db 2370 GAGGGCCCTCCGGCATGTTGCACTCGCTCGCTCTCTGTGAGTGTATGACGAGGCTGT 2429  
QY 501 AlaTyrPyrGluLeuThrProAlaGluThrThrValArgLeuArgAlaTyrMetAsnThr 520  
Db 2430 GCTTGGTATGAGTCAACGCGCGGAGACTACAGTTAGGCTACGAGCGTACATGAACACC 2489  
QY 521 ProGlyLeuProValCysGlnAspHisLeuGluPheTyrPduGlyValPheThrGlyLeu 540  
Db 2490 CCGGGGCTCCCGTGTGCCAGGACCATCTTGAATTTTGGGAGGCGCTCTTTACAGGCGTC 2549  
QY 541 ThrHisIleAspAlaHisPheLeuSerGlnThrLysGlnSerGlyGluAsnLeuProTyr 560  
Db 2550 ACTCATATAGATGCCACTTTCATCCACAAAGCAGAGTGGGAGAACCTTCTTTAC 2609  
QY 561 LeuValAlaTyrGlnAlaThrValCysAlaArgAlaGlnAlaProProSerTrpAsp 580  
Db 2610 CTGGTAGCTACCAAGCCCGGTGTGGCTAGGCTCAAGCCCTCCCTCCATCGTGGAC 2669  
QY 581 GlnMetTrpLysCysLeuIleArgLeuLysProThrLeuHisGlyProThrProLeuLeu 600  
Db 2670 CAGATGTGAAGTGTGTTGATTCGCTCAAGCCACCCCTCCATCGGCGCAACACCCCTGCTA 2729

QY 601 TyrArgLeuGlyAlaValGlnAsnGluValThrLeuThrHisProValThrLysTyrIle 620  
Db 2730 TACAGACTGGCGCTGTTTCCAGAAATGAATCACTCAGCAGCCAGTCCACCAATACATC 2789  
QY 621 MetThrCysMetSerLysAlaAspLeuGluValValThrSerThrTrpValLeuValGlyGly 640  
Db 2790 ATGACATGATGTGCGGCCACCTGGAGGTGCTCACGAGCAGCTGGGTGCTGTGGCGGC 2849  
QY 641 ValLeuAlaAlaLeuAlaAlaTyrCysLeuSerThrGlyCysValValIleValGlyArg 660  
Db 2850 GTCCTGGCTGCTTGGCCCGGATTTGCCCTGTCAACAGGCTGCGTGTGTCATAGTGGCAGG 2909  
QY 661 IleValLeuSerGlyLysProAlaIleIleProAspArgGluValLeuTyrArgGluPhe 680  
Db 2910 GTCGTCTTGTCCGGGAAGCCGCAATCATACCTTGACAGGGAAGTCTCTTACCAGAGTTC 2969  
QY 681 AspGluMetGluCys 686  
Db 2970 GATGAGATGGAAGAGTGC 2987  
RESULT 8  
106440  
LOCUS 106440  
DEFINITION Sequence 54 from Patent EP 0318216.  
ACCESSION 106440  
VERSION 106440.1 GI:590312  
KEYWORDS Unknown.  
SOURCE Unknown.  
ORGANISM Unclassified.  
REFERENCE 1 (bases 1 to 6785)  
AUTHORS Houghton, M., Choo, Q.-L. and Kuo, G.  
TITLE Nanbv diagnostics and vaccines  
JOURNAL Patent: EP 0318216-A1 54 31-MAY-1989;  
FEATURES Location/Qualifiers  
source 1..6785  
/organism="unknown"  
/mol\_type="unassigned DNA"  
ORIGIN  
Alignment Scores:  
Pred. No.: 1.15e-199 Length: 6785  
Score: 3574.00 Matches: 672  
Percent Similarity: 99.42% Conservative: 10  
Best Local Similarity: 97.96% Mismatches: 4  
Query Match: 98.78% Indels: 0  
DB: 6 Gaps: 0  
US-09-930-591-2 (1-6785) x 106440 (1-6785)  
QY 1 MetalaproteinLeuAlaTyrAlaGlnThrArgGlyLeuLeuGlyCysIleIleThr 20  
Db 1203 CTGGCGCCCATCATCGGCGTACGCCAGCAGCAAGGGGCTCTTGGGTGCATAATCACC 1262  
QY 21 SerLeuThrGlyArgAspLysAsnGlnValGluGlyGluValGlnIleValSerThrAla 40  
Db 1263 AGCTTAACCTGGCCGGGACAAAACCAAGTGGAGGGTGAGGTCCAGATTGTCAACTGCT 1322  
QY 41 AlaGlnThrPheLeuAlaThrCysIleAsnGlyValCysTrpThrValTyrHisGlyAla 60  
Db 1323 GCCCAAAACCTTCTTGGCAACGTCATCAATGGGGGTGTGTGGACTGTCTACACGGGGCC 1382  
QY 61 GlyThrArgThrIleAlaSerProLysGlyProValIleGlnMetTyrThrAsnValAsp 80  
Db 1383 GGAACGAGGACCATCGGTCCACCAAGGTCCTGTATCCAGATGATATACCAATGTAGAC 1442  
QY 81 GlnAspLeuValGlyTrpProAlaProGlnGlyAlaArgSerLeuThrProCysThrCys 100  
Db 1443 CAAGACTTGTGGCTGGCCCGCTCCGCAAGGTAGCGCTCATTGACACCTGCACTTGC 1502  
QY 101 GlySerSerAspLeuTyrLeuValThrArgHisAlaAspValIleProValArgArg 120

Db 1503 GGCTCCTCGGACCTTTACCTGGTCAGAGGACGCCGATGTCATTCCTCGTGGCGGGCGG 1562  
Qy 121 GlyAspGlyArgGlySerLeuLeuSerProArgProIleSerTyrLeuLeuGlySerSer 140  
Db 1563 GGTGATAGCAGGGGAGCGCTGCTGTCGCCCGCCCATTTCTCTACTTGAAGGCTCCTCG 1622  
Qy 141 GlyGlyProLeuLeuCysProAlaGlyHisAlaValGlyIlePheArgAlaAlaValCys 160  
Db 1623 GGGGGTCCGCTGTTGTGCGCGCGGGCAGCCGCTCGGCATATTTAGGGCGCGGCTGTGC 1682  
Qy 161 ThrArgGlyValAlaLeuAlaValAspPheIleProValGluSerLeuGluThrMet 180  
Db 1683 ACCCGTGGAGTGGCTAAGCGGTGACATTTATCTCTGTGGAGAACTTAGACAAACCATG 1742  
Qy 181 ArgSerProValPheSerAspAsnSerSerProAlaValProGlnSerTyrGlnVal 200  
Db 1743 AGGTCCCGGTTTACCGGATAACTCTCTCCACCGTAGTGCCCGCAGAGCTTCCAGGTG 1802  
Qy 201 AlaHisLeuHisAlaProThrGlySerGlyLysSerThrLysValProAlaAlaTyrAla 220  
Db 1803 GCTCACTTCCATGCTCCACAGGAGCGGCAAAAGACCAAGAGTCCCGGCTGCATATGCA 1862  
Qy 221 AlaGlnGlyTyrLysValLeuValLeuAsnProSerValAlaAlaThrMetGlyPheGly 240  
Db 1863 GCTCAGGCTATAAGGCTAGTACTCAACCCCTCTGTGTCTGCAACACTGGGCTTTGGT 1922  
Qy 241 AlaTyrMetSerLysAlaHisGlyIleAspProAsnIleArgThrGlyValArgThrIle 260  
Db 1923 GCTTACATGTCAGGCTCATGGATCGATCTTAACATCAGACCGGGGTGAGAACAAAT 1982  
Qy 261 ThrThrGlySerProIleThrTyrSerThrTyrGlyLysPheLeuAlaAspGlyGlyCys 280  
Db 1983 ACCACTGCGACCCCATCACCTACTCCACCTACGCGAAGTTCTTGGCGACCGGGGTGC 2042  
Qy 281 SerGlyAlaTyrAspIleIleCysAspGluCysHisSerThrAspAlaThrSer 300  
Db 2043 TCGGGGGCGCTTATGACATAATAATTTGTACAGAGTGCCACTCCACGGATGCCACATCC 2102  
Qy 301 IleLeuGlyIleGlyThrValLeuAspGlnAlaGluThrAlaGlyAlaArgLeuThrVal 320  
Db 2103 ATCTTGGGCATCGGACTGTCTTGACACAGCAGACTGCGGGGGCGAGACTGTTGTG 2162  
Qy 321 LeuAlaThrAlaThrProGlySerValThrValProHisProAsnIleGluGluVal 340  
Db 2163 CTGCGCACCGCACCCCTCCGGGCTCGCTCACTGTGCCCATCCCAACATCGAGAGGTT 2222  
Qy 341 AlaLeuSerThrThrGlyGluIleProPheTyrGlyLysAlaIleProLeuGluAlaIle 360  
Db 2223 GCTCTGTCCACCGGAGAGATCCCTTTTACCGGCAAGGCTATCCCTCTCGAAGTAATC 2282  
Qy 361 LysGlyGlyArgHisLeuIlePheCysHisSerLysLysCysAspGluLeuAlaAla 380  
Db 2283 AAGGGGGGAGACATCTCTCTGTCTTCAAGAGAGAGTGGCGAGAACTCGCGCA 2342  
Qy 381 LysLeuValAlaLeuGlyValAsnAlaValAlaTyrTyrArgGlyLeuAspValSerVal 400  
Db 2343 AAGCTGTGTGATGGCATCAATCGCGTCTACTACCGCGTCTTTCAGTGTCCGTC 2402  
Qy 401 IleProThrSerGlyAspValValAlaThrAspAlaLeuMetThrGlyPheThr 420  
Db 2403 ATCCCGACCGCGCGATGTTGTGTCGTCGCAACCGATGCCCTCATGACCGGTATACC 2462  
Qy 421 GlyAspPheAspSerValIleAspCysAsnThrCysValThrGlnThrValAspPheSer 440  
Db 2463 GGCAGCTTCGACTCGGTGATAGATGCAATACGTGTGTACCCAGACAGTCGATTTCCAGC 2522  
Qy 441 LeuAspProThrPheThrIleGluThrIleThrLeuProGlnAspAlaValSerArgThr 460  
Db 2523 CTTCACCTACTTCCACATTTGACACATCAGCTCCCGCAGGATGTGTCTCCCGACT 2582  
Qy 461 GlnArgArgGlyArgThrGlyArgGlyLysProGlyIleTyrArgPheValAlaProGly 480  
Db 2583 CAACGTCGGGCGAGGACTGGCAGGGGGAAGCCAGGCATCTACAGATTTGTGGCACCGGGG 2642

Qy 481 GluArgProSerGlyMetPheAspSerSerValLeuCysGluCysTyrAspAlaGlyCys 500  
Db 2643 GAGCGCCCTCCGCATGTTCCGACTCGTCCGTCTCTGTGAGTGTATGACGAGGCTGT 2702  
Qy 501 AlaTyrTyrGluLeuThrProAlaGluThrThrValArgLeuArgAlaTyrMetAsnThr 520  
Db 2703 GCTTGGTATGAGCTCACGCCCGCGAGACTACAGTTAGGCTACGAGCGTACATGAACACC 2762  
Qy 521 ProGlyLeuProValCysGlnAspHisLeuGluPheTyrGluGlyValPheThrGlyLeu 540  
Db 2763 CCGGGGGTTCCTGTTGTCAGGACCACTTGAATTTTGGAGGGCGTCTTTACAGGCCCTC 2822  
Qy 541 ThrHisIleAspAlaHisPheLeuSerGlnThrLysGlnSerGlyGluAsnLeuProTyr 560  
Db 2823 ACTCATATAGATGCCACTTCTATCCAGACAAAGCAGAGTGGGGAGAACCTTCTTAC 2882  
Qy 561 LeuValAlaTyrGlnAlaThrValCysAlaArgAlaGlnAlaProProSerTyrPasp 580  
Db 2883 CTGGTAGCTACCAAGCCACCGTGTGGCTAGGCTCAAGCCCTCCCCCATCGTGGAC 2942  
Qy 581 GlnMetTyrLysCysLeuIleArgLeuLysProThrLeuHisGlyProThrProLeuLeu 600  
Db 2943 CAGATGTGAAGTGTGATTCGCCTCAAGCCCACTTCCATGGGCCAACCCCTTGCTA 3002  
Qy 601 TyrArgLeuGlyAlaValGlnAsnGluValThrLeuThrHisProValThrLysTyrIle 620  
Db 3003 TACAGACTGGGGCTGTTCAGATGAATCAACCTGACGACCCAGTCACCAATACATC 3062  
Qy 621 MetThrCysMetSerAlaAspLeuGluValValThrSerThrTyrValLeuValGlyGly 640  
Db 3063 ATGACATGATGTCCGCCGACCTGGAGTGTCTACGAGCACCTGGGTGTCTGTTGGCGGC 3122  
Qy 641 ValLeuAlaAlaLeuAlaTyrCysLeuSerThrGlyCysValValIleValIleArg 660  
Db 3123 GTCTGTGCTGTTTGGCGCGGTATGCTGTCAACAGGCTGCGGTGCATAGTGGGCGAG 3182  
Qy 661 IleValLeuSerGlyLysProAlaIleIleProAspArgGluValLeuTyrArgGluPhe 680  
Db 3183 GTGCTTGTTCGGGAGCGGCATCATCTGACAGGGNAGTCTTACCGAGAGTTC 3242  
Qy 681 AspGluMetGluGluCys 686  
Db 3243 GATGAGATGAAGAGTGC 3260

## RESULT 9

I09329  
LOCUS I09329 6785 bp DNA linear PAT 02-DEC-1994  
DEFINITION Sequence 10 from Patent WO 8904669.  
ACCESSION I09329  
VERSION I09329.1 GI:587964  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unclassified.  
REFERENCE 1 (bases 1 to 6785)  
AUTHORS Houghton, M., Choo, Q.-K. and Kuo, G.  
JOURNAL Patent: WO 8904669-A 10 01-JUN-1989;  
FEATURES  
Location/Qualifiers  
1..6785  
/organism="unknown"  
/mol\_type="unassigned DNA"

## ORIGIN

Alignment Scores:  
Pred. No.: 1.15e-199 Length: 6785  
Score: 3574.00 Matches: 672  
Percent Similarity: 99.42% Conservative: 10  
Best Local Similarity: 97.96% Mismatches: 4  
Query Match: 98.78% Indels: 0  
DB: 6 Gaps: 0

US-09-930-591-2 (1-686) x I09329 (1-6785)

```
QY 1 MetAlaProIleThrAlaTyrAlaGlnGlnThrArgGlyLeuLeuGlyCysIleIleThr 20
Db : : : : :
1203 CTGGCGCCCATCACGGGTACGCCAGCAGACAAAGGGGCTCTCGGTGCATAATCAC 1262

QY 21 SerLeuThrGlyArgAspLysAenGlnValGluGlyGluValGlnIleValSerThrAla 40
Db : : : : :
1263 AGCTAACTGGCGGGGACAAAACCAAGTGGAGGGTGAGGTCCAGATTGTGTCAACTGCT 1322

QY 41 AlaGlnThrPheLeuAlaThrCysIleAenGlyValCysTrpThrValTyrHisGlyAla 60
Db : : : : :
1323 GCCCAACCTTCTCGCAACGTGCATCAATGGGGGTGTCTGGACTGTCTACCAAGGGGCC 1382

QY 61 GlyThrArgThrIleAlaSerProLysGlyProValIleGlnMetTyrThrAenValAsp 80
Db : : : : :
1383 GGNACGAGGACCATCGGTCACCAAGGGTCTGTCTCATCCAGATGATACCAATGTAGAC 1442

QY 81 GlnAspLeuValGlyTrpProAlaProGlnGlyAlaArgSerLeuThrProCysThrCys 100
Db : : : : :
1443 CAAGACTTGTGGCTGGCGCGCTCCGCAAGGTAGCGGCTCAATTGACACCTGCACCTGC 1502

QY 101 GlySerSerAspLeuTyrIleuValThrArgHisAlaAspValIleProValArgArgArg 120
Db : : : : :
1503 GGCTCTCTCGGACCTTACCTGGTTCAGAGCACGCGGATGTCATCCGGTGGCGCGGG 1562

QY 121 GlyAspGlyArgGlySerLeuLeuSerProArgProIleSerTyrLeuLysGlySerSer 140
Db : : : : :
1563 GGTGATAGAGGGGACGCTGTCTGCCCGCGGCCCATTTCTTCTTGAAGGCTCTCTCG 1622

QY 141 GlyGlyProLeuLeuCysProAlaGlyHisAlaValGlyIlePheArgAlaAlaValCys 160
Db : : : : :
1623 GGGGGTCCGCTGTGTGCCCGCGGGGACCGCTGGGCATATTTAGGGCGCGGGTGTGC 1682

QY 161 ThrArgGlyValAlaLysAlaValAspPheIleProValGlnSerLeuGluThrThrMet 180
Db : : : : :
1683 ACCCGTGGAGTGGCTAAGCGGGTGGACTTTATCCCTGTGGAGAACCTTAGAGACAAACATG 1742

QY 181 ArgSerProValPheSerAspAenSerSerProProAlaValProGlnSerTyrGlnVal 200
Db : : : : :
1743 AGTCTCCCGGTGTTCAGGATTAATCTCTCTCCACGAGTAGTCCCGACAGCTTCCAGGTG 1802

QY 201 AlaHisLeuHisAlaProThrGlySerGlyLysSerThrLysValProAlaAlaTyrAla 220
Db : : : : :
1803 GCTCACTCCATGCTCCACAGCGCAGCGGCAAAAGCACCAAGTCCCGGTGCATATGCA 1862

QY 221 AlaGlnGlyTyrLysValIleuValLeuAenProSerValAlaAlaThrMetGlyPheGly 240
Db : : : : :
1863 GCTCAGGGCTATTAAGGTGTGTAGTACTCAACCCCTCTGTGCTGCAACACTGGGGCTTGGT 1922

QY 241 AlaTyrMetSerLysAlaHisGlyIleAspProAsnIleArgThrGlyValArgThrIle 260
Db : : : : :
1923 GCTTACATGTCGAAGGCTCATGGGATCGATCTTAACATCAGGACCGGGGTGAGAACAT 1982

QY 261 ThrThrGlySerProIleThrTyrSerThrTyrGlyLysPheLeuAlaAspGlyGlyCys 280
Db : : : : :
1983 ACCACTGGGACGCCCATCAGTACTCCACCTACGGCAAGTTCCTTGGCCGAGCGGGTGC 2042

QY 281 SerGlyGlyAlaTyrAspIleIleCysAspGluCysHisSerThrAspAlaThrSer 300
Db : : : : :
2043 TCGGGGGGCGCTTATGACATAATAATTTGTGAGGAGTGGCCACTCCACGGATGCCACATCC 2102

QY 301 IleLeuGlyIleGlyThrValLeuAspGlnAlaGluThrAlaGlyAlaArgLeuThrVal 320
Db : : : : :
2103 ATCTTGGGCATCGGCACTGTCTTGACCAAGCAGAGACTCGGGGGGCGAGCTGGTGTG 2162

QY 321 LeuAlaThrAlaThrProGlySerValThrValProHisProAenIleGluGluVal 340
Db : : : : :
2163 CTGCCACCGCCACCCCTCGGGCTCGGTCACTGTGCCCATCCCAACATCGAGGAGTT 2222

QY 341 AlaLeuSerThrThrGlyGluIleProPheTyrGlyLysAlaIleProLeuGluAlaIle 360
Db : : : : :
2223 GCTCTGTGCCACCGGAGAGATCCCTTTTACGGCAAGGCTATCCCTTCCGAGTAAATC 2282
```

## RESULT 10

```
LOCUS AR118696 7310 bp DNA linear PAT 16-MAY-2001
DEFINITION Sequence 74 from patent US 6150087.
ACCESSION AR118696
```

```
QY 361 LysGlyGlyArgHisLeuIlePheCysHisSerLysLysCysAspGluLeuAlaAla 380
Db : : : : :
2283 AAGGGGGGCGAGACATCTCATCTCTCTCTCAAGAGAGAGTGCAGCACTCGCGCA 2342

QY 381 LysLeuValAlaLeuGlyValAsnAlaValAlaTyrTyrArgGlyLeuAspValSerVal 400
Db : : : : :
2343 AAGCTGGTGCATTTGGGCATCAATCCGTGGCTACTACCGGGTCTTGACGTTCGTC 2402

QY 401 IleProThrSerGlyAspValValValAlaAlaThrAspAlaLeuMetThrGlyPheThr 420
Db : : : : :
2403 ATCCCGACGACGCGGATGTTGTCTGTGGCAACCGATGCCCTCATGCCGCTATACC 2462

QY 421 GlyAspPheAspSerValIleAspCysAenThrCysValThrGlnThrValAspPheSer 440
Db : : : : :
2463 GCGCACTTCGACTCGGTGATAGACTGCAATACGTGTGTCCACGACAGTCGATTTCCAG 2522

QY 441 LeuAspProThrPheThrIleGluThrIleThrLeuProGlnAspAlaValSerArgThr 460
Db : : : : :
2523 CTGTGACCTTACCTTTCACCATGAGACATCAGCTCCCCAGGATGCTGTCTCCGCACT 2582

QY 461 GlnArgArgGlyArgThrGlyArgGlyLysProGlyIleTyrArgPheValAlaProGly 480
Db : : : : :
2583 CAACGTGCGGGCAGGACTGCGAGGGGGAAGCCAGGCATCTACAGATTGTGGCACCGGG 2642

QY 481 GluArgProSerGlyMetPheAspSerSerValLeuCysGluCysTyrAspAlaGlyCys 500
Db : : : : :
2643 GAGCGCCCTCCGGCATGTTCCGACTCGTCCGTCTCTGTGAGTGTATGACGCGAGGTGT 2702

QY 501 AlaTrpTyrGluLeuThrProAlaGluThrThrValArgLeuArgAlaTyrMetAenThr 520
Db : : : : :
2703 GCTTGTGTATGAGCTACGCCCGCGGAGACTACAGTTAGGCTACGAGGTACATGACACC 2762

QY 521 ProGlyLeuProValCysGlnAspHisLeuGluPheTrpGluGlyValPheThrGlyLeu 540
Db : : : : :
2763 CCGGGGCTTCCCGTGTGCCAGGACCATCTTGAATTTGGAGGGCGCTTTTACAGGCTC 2822

QY 541 ThrHisIleAspAlaHisPheLeuSerGlnThrLysGlnSerGlyGluAenLeuProTyr 560
Db : : : : :
2823 ACTCATATAGATGCCCACTTCTATCCAGACAAAGCAGAGTGGGAGAACCTTCTCTTAC 2882

QY 561 LeuValAlaTyrGlnAlaThrValCysAlaArgAlaGlnAlaProProSerTrpAsp 580
Db : : : : :
2883 CTGGTAGCTGATACCAAGCCCGGTGTGGCTAGGGCTCAAGCCCTCCCCCATCGTGGGAC 2942

QY 581 GlnMetTrpLysCysLeuIleArgLeuLysProThrLeuHisGlyProThrProLeuLeu 600
Db : : : : :
2943 CAGATGTGGAAGTGTGTGATTCCCTCAAGCCCAACCTCCATGGGCCAACCCCTGCTA 3002

QY 601 TyrArgLeuGlyAlaValGlnAenGluValThrLeuThrHisProValThrLysTyrIle 620
Db : : : : :
3003 TACAGACTGGCGCTGTTCAGAAATGAAATCACCCCTGACGCCACCCAGTCACCAATACATC 3062

QY 621 MetThrCysMetSerAlaAspLeuGluValValThrSerThrTrpValLeuValGlyGly 640
Db : : : : :
3063 ATGACATGATATGTGCGGCCACCTGGAGGTGCTCAGGACACCTGGGGTCTCGTTGGCGG 3122

QY 641 ValLeuAlaAlaLeuAlaAlaTyrCysLeuSerThrGlyCysValValIleValGlyArg 660
Db : : : : :
3123 GTCTGGCTGTCTTGGCGGGTATTCCTGTCAACAGGCTGCGGTGTCATAGTGGGCGAG 3182

QY 661 IleValLeuSerGlyLysProAlaIleIleProAspArgGluValLeuTyrArgGluPhe 680
Db : : : : :
3183 GTCGTCTGTGTCGGGAAGCGGCAATCATACCTGACAGGGAAGTCTCTTACCGAGAGTTC 3242

QY 681 AspGluMetGluClnCys 686
Db : : : : :
3243 GATGAGATGGAGAGGTGC 3260
```

VERSION	AR118696.1	GI:14100606
KEYWORDS	Unknown.	
SOURCE	Unknown.	
ORGANISM	Unclassified.	
REFERENCE	1 (bases 1 to 7310)	
AUTHORS	Chien,D.Y.	
TITLE	NANBV diagnostics and vaccines	
JOURNAL	Patent: US 6150087-A 74 21-NOV-2000;	
FEATURES	Location/Qualifiers	
source	1..7310	
	/organism="unknown"	
ORIGIN	/mol_type="unassigned DNA"	
Alignment Scores:		
Pred. No.:	1.26e-199	Length: 7310
Score:	3574.00	Matches: 672
Percent Similarity:	99.42%	Conservative: 10
Best Local Similarity:	97.96%	Mismatches: 4
Query Match:	98.78%	Indels: 0
DB:	6	Gaps: 0
US-09-930-591-2 (1-686) x AR118696 (1-7310)		
Qy	1	MetAlaProIleThrAlaTy rAlaGlnGlnThrArgGlyLeuLeuGlyCysIleIleThr 20
Db	1728	CTGGCGGCCATCAGCGGTACGCCAGCAGACAAGGGGCTCTCTAGGGTGCATAATCACC 1787
Qy	21	SerLeuThrGlyArgAspLysAsnGlnValGluGlyGluValGlnIleValSerThrAla 40
Db	1788	AGCTTAAGTGGCGGGCAAAAACCAAGTGGAGGGTGAGGTCCAGATTGTCAACTGCT 1847
Qy	41	AlaGlnThrPheLeuAlaThrCysIleAsnGlyValCysTrpThrValTyrHisGlyAla 60
Db	1848	GCCCAAACTTCCTGGCAACGTGCATCAATGGGGTGTGTGGACTGTCTACCACGGGCC 1907
Qy	61	GlyThrArgThrIleAlaSerProLysGlyProValIleGlnMetTyrThrAsnValAsp 80
Db	1908	GGAAAGGAGGACATCGCGTCAACCAGAGGGTCTGTCTATCCAGATGTATACCAATGTAGAC 1967
Qy	81	GlnAspLeuValGlyTyrProAlaProGlnGlnAlaArgSerLeuThrProCysThrCys 100
Db	1968	CAAGACCTTGTGGGTGGCCCGCTCCGCAAGGTAGCGCTCATTGACACCTTGCACTTGC 2027
Qy	101	GlySerSerAspLeuTyrLeuValThrArgHisAlaAspValIleProValArgArg 120
Db	2028	GGGTCTCTCGGACCTTTACCTGGTCACGAGGCACCGCATGTCAATCCGTGGCGCGCGG 2087
Qy	121	GlyAspGlyArgGlySerLeuSerProArgProIleSerTyrLeuLysGlySerSer 140
Db	2088	GGTGATAGAGGGGACGCTGCTGCGCCCGGCCCATTTCTACTTTGAAGGGTCTCTCG 2147
Qy	141	GlyGlyProLeuLeuCysProAlaGlyHisAlaValGlyIlePheArgAlaAlaValCys 160
Db	2148	GGGGGTCCGCTGTTGTGCCCCGGGGGCACGCGTGGGCATATTTAGGCGCGGTGTGC 2207
Qy	161	ThrArgGlyValAlaLysAlaValAspPheIleProValGluSerLeuGluThrThrMet 180
Db	2208	ACCGTGGAGTGGCTAAGGGCGTGGACTTTATCCCTGTGGAGAACCTTAGAGACAACCATG 2267
Qy	181	ArgSerProValPheSerAspAsnSerSerProProAlaValProGlnSerTyrGlnVal 200
Db	2268	AGGTCCCGGTGTTCCGGATACTCTCTTCACCAAGTAGTGCCCAAGAGCTTCCAGGTG 2327
Qy	201	AlaHisLeuHisAlaProThrGlySerGlyLysSerThrLysValProAlaAlaTyrAla 220
Db	2328	GCTCACCTCCATGCTCCCAAGGCAGCGGCAAAAGCACCAAGGTCCCGGTGCATATGCA 2387
Qy	221	AlaGlnGlyTyrLysValLeuValLeuAsnProSerValAlaAlaThrMetGlyPheGly 240
Db	2388	GCTCAGGGCTATAGGTGTGTAAGTACTCAACCCCTCTGTGTGTCGAACACTGGGTGGT 2447



```
Db 3528 TACAGATGGCGCTGTTTCAAGTAAATCACCCTGACGACCCAGTCCACCAATACATC 3587
Qy 621 MetThrCysMetSerAlaAspLeuValValThrSerThrTrpValLeuValGlyGly 640
Db 3588 ATGACATGCATGTCGGCGGACCTGGAGGTGCTCACGAGCACCTGGGTGGTGGCGGC 3647
Qy 641 ValLeuAlaLeuAlaAlaTyrCysLeuSerThrGlyCysValValIleValGlyArg 660
Db 3648 GTCTGTGCTGTTTGGCGCGGTATGCTGCTCAACAGCTGCGGTGTCATAGTGGCAGG 3707
Qy 661 IleValLeuSerGlyLysProAlaIleIleProAspArgGluValLeuTyrArgGluPhe 680
Db 3708 GTGCTGTTGTCGGGAAGCGGCAATCATACCTGACAGGGAAGTCTCTACCGAGAGTTC 3767
Qy 681 AspGluMetGluGluCys 686
Db 3768 GATGAGATGGAAGAGTGC 3785

RESULT 11
LOCUS I09331 7310 bp DNA linear PAT 02-DEC-1994
DEFINITION Sequence 15 from Patent WO 8904669.
ACCESSION I09331
VERSION I09331.1 GI:587966
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 7310)
AUTHORS Houghton,M., Choo,Q.-K. and Kuo,G.
JOURNAL Patent: WO 8904669-A 15 01-JUN-1989;
FEATURES
    Location/Qualifiers
        source
            1..7310
                /organism="unknown"
                /mol_type="unassigned DNA"

ORIGIN

Alignment Scores:
Pred. No.: 1..26e-199 Length: 7310
Score: 3574.00 Matches: 672
Percent Similarity: 99.42% Conservative: 10
Best Local Similarity: 97.96% Mismatches: 4
Query Match: 98.78% Indels: 0
DB: 6 Gaps: 0

US-09-930-591-2 (1-686) x I09331 (1-7310)
Qy 1 MetAlaProIleThrAlaTyrAlaGlnGlnThrArgGlyLeuLeuGlyCysIleIleThr 20
Db 1728 CTGGCGCCCATCACGGGCTAGCCGACAGACAAAGGGCCCTCTAGGGTGCATANTACC 1787
Qy 21 SerLeuThrGlyArgAspLysAsnGlnValGluGlyGluValGlnIleValSerThrAla 40
Db 1788 AGCCTAACTGCGCGGGACAAAACCAAGTAGGAGGTGAGGTCCAGATTGTCAACTGCT 1847
Qy 41 AlaGlnThrPheLeuAlaThrCysIleAsnGlyValCysTrpThrValTyrHisGlyAla 60
Db 1848 GCCCAACCTTCTCGGACAGTGCATCAATGGGGTGTGCGGACTGTCTACACAGGGGCC 1907
Qy 61 GlyThrArgThrIleAlaSerProLysGlyProValIleGlnMetTyrThrAsnValAsp 80
Db 1908 GGAACGAGGACCATCGGCTCACCAAGGGTCTGTCTATCCAGATGTATACCAATGTAGAC 1967
Qy 81 GlnAspLeuValGlyTrpProAlaProGlnGlyAlaArgSerLeuThrProCysThrCys 100
Db 1968 CAAGACCTTGTGGCTGGCCGCTCGCAAGAGTAGCGGCTCAATTGACACCCCTGCATTGC 2027
Qy 101 GlySerSerAspLeuTyrLeuValThrArgHisAlaAspValIleProValArgArgArg 120
Db 2028 GGCTCTCGGACCTTTACCTGGTTCACGAGCAGCCGATGTCAATCCGTGGCCGCGCGG 2087
Qy 121 GlyAspGlyArgGlySerLeuLeuSerProArgProIleSerTyrLeuLysGlySerSer 140
```

```
Db 2088 GGTTAGTACGAGGGGAGCTGTGTGTCGCCCGGCCCATTTCTTCTACTTTGAAAGGCTCCTCG 2147
Qy 141 GlyGlyProLeuLeuCysProAlaGlyHisAlaValGlyIlePheArgAlaAlaValCys 160
Db 2148 GGGGGTCCGCTGTGTGTCGCCCGGGGACGCGCGTGGGCATATTTAGGGCGGGGTGTGC 2207
Qy 161 ThrArgGlyValAlaLysAlaValAspPheIleProValGluSerLeuGluThrThrMet 180
Db 2208 ACCGTGGAGTGGCTAAGCGGTGGACTTTATCCCTGTGGAGAACCTTAGAGACAACCATG 2267
Qy 181 ArgSerProValPheSerAspAsnSerSerProAlaValProGlnSerTyrGlnVal 200
Db 2268 AGTCCCCCGGTGTTCCGGATNACTCTCTCCACCATAGTAGTCCCGAGCTTCCAGGTG 2327
Qy 201 AlaHisLeuHisAlaProThrGlySerGlyLysSerThrLysValProAlaAlaTyrAla 220
Db 2328 GCTCAGCTCCATGCTCCACAGGACGCGCAAGCAACCAAGGTCCCGGTGCATATGCA 2387
Qy 221 AlaGlnGlyTyrLysValLeuValLeuAsnProSerValAlaAlaThrMetGlyPheGly 240
Db 2388 GCTCAGGGCTATAGGTGTAGTACTCAACCCCTCTGTTGCTGCAACACACTGGGCTTTGGT 2447
Qy 241 AlaTyrMetSerLysAlaHisGlyIleAspProAsnIleArgThrGlyValArgThrIle 260
Db 2448 GCTTACATGTCCAAGGCTCATGGATCGATCCTTAACATCAGGACCGGGGTGAGAACAAAT 2507
Qy 261 ThrThrGlySerProIleThrTyrSerThrTyrGlyLysPheLeuAlaAspGlyGlyCys 280
Db 2508 ACCATGGCAGCCCATCATCGTACTCCACCTACGCAAGTTCCTTCCGACGCGGGTGC 2567
Qy 281 SerGlyGlyAlaTyrAspIleIleCysAspGluCysHisSerThrAspAlaThrSer 300
Db 2568 TCGGGGGCGCTTATGACATATATTTGTGACAGTGCACCTCCACGGATGCCACATCC 2627
Qy 301 IleLeuGlyIleGlyThrValLeuAspGlnAlaGluThrAlaGlyAlaArgLeuThrVal 320
Db 2628 ATCTTGGCATCGGCATGTCTTACCAAGCAGAGACTCGCGGGCGGAGACTGGTTGTG 2687
Qy 321 LeuAlaThrAlaThrProGlySerValThrValProHisProAsnIleGluGluVal 340
Db 2688 CTCGCCACCGCCACCCCTCGGCTCCGTCACCTGTCGCCCATCCCAACATCGAGGAGTT 2747
Qy 341 AlaLeuSerThrThrGlyGluIleProPheTyrGlyLysAlaIleProLeuGluAlaIle 360
Db 2748 GCTCTGTCCACCACCGAGAGATCCCTTTTACGCGCAAGGCTATCCCTTCGAAGTAATC 2807
Qy 361 LysGlyGlyArgHisLeuIlePheCysHisSerLysLysCysAspGluLeuAlaAla 380
Db 2808 AAGGGGGGAGACATCTCATCTTCTGTCTTCAAAGAAGAGTGCAGCAACTCGCGCA 2867
Qy 381 LysLeuValAlaLeuGlyValAsnAlaValAlaTyrTyrArgGlyLeuAspValSerVal 400
Db 2868 AAGCTGTGCGCATTTGGGCTCAATGCGTGGCTACTACCGGGTCTTGACGTCGTCGTC 2927
Qy 401 IleProThrSerGlyAspValValValValAlaThrAspAlaLeuMetThrGlyPheThr 420
Db 2928 ATCCCGACGAGCGCGATGTTGTCGTCGTCGCGAACCGATGCCCTCATGCCGCTATACC 2987
Qy 421 GlyAspPheAspSerValIleAspCysAsnThrCysValThrGlnThrValAspPheSer 440
Db 2988 GCGCATTCGACTCGGTGATAGACTGCAATACGTGTGTCAACCAGACAGTCGATTTTCAGC 3047
Qy 441 LeuAspProThrPheThrIleGluThrIleThrLeuProGlnAspAlaValSerArgThr 460
Db 3048 CTTGACCCCTACCTTCACTTACCAATGAGCAATCAGCTCCCGCAGGATGCTGTCTCCGCACT 3107
Qy 461 GlnArgArgGlyArgThrGlyArgGlyLysProGlyIleTyrArgPheValAlaProGly 480
Db 3108 CAACTGCGGGCAGGACTGGCAGGGGAGCCAGGCGATCTACAGATTTGTGGCACCGGGG 3167
Qy 481 GluArgProSerGlyMetPheAspSerValLeuCysGluCysTyrThrAspAlaGlyCys 500
```



Db	3168	GAGCGCCCTCGGCATGTTTCAGACTGCTCGCTCTCTGTGAGTGTATGACGCGAGGTGT	3227
Qy	501	AlaTrpTyrGluLeuThrProAlaGluThrThrValArgLeuArgAlaTyrMetAsnThr	520
Db	3228	GCTTGGTATGAGTTCAGCGCCCGCAGAGACTACAGTTAGCTACGAGCGTACATGAACACC	3287
Qy	521	ProGlyLeuProValCysGlnAspHisLeuGluPheTrpGluGlyValPheThrGlyLeu	540
Db	3288	CCGGCGCTCCCGTGTGCGAGACCATCTTGAATTTTGGAGGGCGCTCTTTACAGCGCTC	3347
Qy	541	ThrHisTleAspAlaHisPheLeuSerGlnThrLysGlnSerGlyGluAsnLeuProTyr	560
Db	3348	ACTCATATAGATGCCACTTTCTATCCAGAACAGAGAGTGGGAGAACCTTCTCTTAC	3407
Qy	561	LeuValAlaTyrGlnAlaThrValCysAlaArgAlaGlnAlaProProProSerTrpAsp	580
Db	3408	CTGGTAGCTACCAAGCCACCGTGTGCGCTAGGGCTCAAGCCCTCCGCCATCGTGGAC	3467
Qy	581	GlnMetTrpLysCysLeuIleArgLeuLysProThrLeuHisGlyProThrProLeuLeu	600
Db	3468	CAGATGTGGAAGTGTTCATTCGCTCAAGCCACCTCCATCGGCGCAACACCCCTGCTA	3527
Qy	601	TyrArgLeuGlyAlaValGlnAsnGluValThrLeuThrHisProValThrLysTyrIle	620
Db	3528	TACAGATGGCGCGTGTTCAGAAATCAACCTCGACGCGCCAGTCCACCAATACATC	3587
Qy	621	MetThrCysMetSerAlaAspLeuGluValValThrSerThrTrpValLeuValGlyGly	640
Db	3588	ATGACATGATGTCGGCGCAGCTGGAGTCTGTCAGGACACTGGGTGCTGTTGGCGGC	3647
Qy	641	ValLeuAlaAlaLeuAlaAlaTyrCysLeuSerThrGlyCysValValIleValGlyArg	660
Db	3648	GTCCTGGCTGTTTGGCGCGCTATTGCTGTCACAGCGCTCGTGGTTCATAGTGGCAGG	3707
Qy	661	IleValLeuSerGlyLysProAlaIleIleProAspArgGluValLeuTyrArgGluPhe	680
Db	3708	GTCCTGTGTCGGGAGCGGCAATACATCTGACAGGGAAGTCTCTACCGAGAGTTC	3767
Qy	681	AspGluMetGluGluCys 686	
Db	3768	GATGAGATGGAGAGTGC 3785	
RESULT 12			
HPCPOLYP			
LOCUS	HPCPOLYP	7310 bp ss-RNA	linear
DEFINITION	Hepatitis C virus polyprotein gene, partial cds.		VRL 02-AUG-1993
ACCESSION	M32084		
VERSION	M32084.1	GI:329875	
KEYWORDS	polyprotein.		
SOURCE	Hepatitis C virus		
ORGANISM	Hepatitis C virus		
REFERENCE	1	(bases 1 to 7310)	
AUTHORS	Choo, Q.-L., Richman, K. and Han, J.		
TITLE	The nucleotide sequence of the Hepatitis C viral genome		
JOURNAL	Unpublished (1990)		
COMMENT	Original source text: Hepatitis C virus, cDNA to viral RNA, clones K9-1 through 15e, isolated from chimpanzee (individual 910) blood plasma. Draft entry and printed sequence for [1] kindly submitted by M.Houghton, 22-FEB-1990. Chiron Corporation, 4560 Horton Street, Emeryville, CA 94608.		
FEATURES	Location/Qualifiers		
source	1..7310		
	/organism="Hepatitis C virus"		
	/mol_type="genomic RNA"		
	/db_xref="taxon:11103"		
	<1..>7310		
	/note="polyprotein"		
	/codon_start=3		
CDS	/protein_id="AAA45677.1"		
	/db_xref="GI:329876"		

/translation="GCPERLASCRPLTDFDQGWGPISYANGSGDQDRPYCWHYPPKPC  
 GIVPAKSCVPCVYCTPSPVVTGDRSGAPTYSGENDTDVFLNLTNPTGNGWFC  
 TWNSGTFTKVGAPPCVIGAGNTLHCTDPCRKHPTDATYSCRGSPWITPRCLVD  
 YPRVLMHVPCTINVTIPIKIMYGVGVHRELEAACNWRBRCRDLDRSLSPLLLT  
 TQOMVLPCTFTTLPALSTGLIHLQNIIVDVOYLXGVGSSIASWAIKEWYVLLPLLL  
 ADARVCSCLNMMLLSIQEALENLVILNAAASLAGTHGLVFLVFFCFANLYLKKWVP  
 GAVTYFGMPPLLLLLLALPQRAYLDEVAASCGVVLVGLMALTISPPYKRYISWC  
 LMWLYFLTRVEAQLHVMIPPLNVRGGRDAVILLMCAVHTPLVFDITKLLAVLPGFLW  
 ILQASLLKVPYFVRVQGLLRFCALARMKGGHYQVMVILKGLATGYVYVNHLLTPAD  
 WAHNGRLDLAVAVBPVFSOMETKLIITWGDATACGDIILNGLVSARRRGRLGPAD  
 GMVSKGWRLLAPITAYAOQTRGLGCIITSLTGRDKMQVEGEVQIVSTAQTFLATCI  
 NGVCWTYHGAGTRTIAAPKGPVQMTINVDQDLVGPAPQSRSLTPTCTCGSSLDY  
 TRHADVIPVRRGDSGLSPRPISYLKSGSGPLCPAGHAGVTFPRAAVCTRGVA  
 KAVDIFIVENLETTRMSPVFNDSPPVQSFQVFAHLHAPTSGKSTKVPAAVAQG  
 KYVLNLPNSVAATLFGAYMSKAGHIDPNITGVRTITGSPITVSTYSGKLDGCGS  
 GGAVDIIICDBCHSTDATSLIGICTVLDOAETAGARLVLATATPPGSVTPHPNIE  
 VALSTGEIIPYKAIIPLEVIKGRHLIFCHSKKCBELAKALVALGINAVAYRGLD  
 VSVIPTSGDVVVAITDMTGYTGFDSVIDCNCTVCTVDFTSLDPTFTITLTPQD  
 AVSRQTRGRTRGKPGIYRFVAFGRPSGMFDSVLCCEYDAGCAMYELTPAETVR  
 LRAYMNTPGLPVQDHLFEWEGFTGLTHDAHLSQTKSGENLPYLVAQYATCAR  
 AQAAPPDMDWMKCLIRLPTLHGPTPLLYRLGAVONEITLTHPVKYIMTCSADLE  
 VVTSVLVGVLAALAAAYCLSTGCVVIIGRVVLSGKPAITPDREVLRYREDEMBECS  
 QHLPYIQGMMLEAQFKALGLLOTAQRQAEVIAPAVOTWQKLETFWAKMNFIS  
 GIQYLAGSLTPGNPAIASLMAFTAAVTSPLTTSQTLFLNLGGWVAQAALAPGAATA  
 FVAGLAGAAGISVGLKVLIDILAGYGAGVAGALVAFKIMSGSPVSTEDLVNLLPAI  
 LSPCALVGVVVCRAILRRHVGPGEQVQMMNRLIAFASRGNHVSPTHVPESDAAARV  
 TAILSSLTVQLRLRLHOWISSECTTPCSGSLRDINDWICEVLSDFKWLKAKLMPQ  
 LPGIPVSCORGKYGWVRDGIHMTIRCHCGAEITGHVKNGTMRIVGPRTCNMWSGTF  
 PINATYTCPTPLPAPNYTALWRVSAEVEIRQVGDHYVTCPTDNLNKKPCQVPS  
 PERFTELDGVRLLHFAAPPKTLREESVFRVGLHEVPVGSQPCPEPDDVVALYSMTL  
 QHETIAAARGLARGSPSVASSASQASPSLTKATCTANHDSPOAELTEANLWR  
 QMGENTIRVSEKNVILDSFOPLVAEEDEREISVPAEILKRRRAQALPFWARPD  
 YNPFLVETKPKDPYEPVPHGCPPLPPKPPVPPKRTVLTTESTLSTALAEALT  
 SFSSSTSGITGDNNTTSSPAPGCPDSDAESYSSMPLEGBGPDLDLSGKSTVR  
 SSANADVVCCSMYSWTGALVTPCAAEQKLPINALSNLLRHHLLVYSTSRAC  
 QROKQVTRDLQVLDVSHQVLEKAAASKVANLLSVEACSLTPPHSKAKPGYV  
 AKVDRCHARKAVTHINSVWKLLEDNTVIDTITIMAKNEVFCVQPGRGRKAPKLVF  
 PDGLRVYCEKMALYDVVTKLPVAMVSSYGFQSPQGVQVFLVQAMKSKPTPMGFYR  
 TRCPDSTVTBESDIRTEBAIQOCDLPOARVAIKSLTERLVVGLPLNTRGECGYRR  
 CRASGLVLTSCGNTLTCYIKARACRAAGLQDCTMLVCGDDLVVICBSAGVQEDAAL  
 RAFTAMTRYSAAPPDPPQPEYDELITSCSSNVSVVAHDGKGVYILTRDPTTFLAR  
 AAWEATRHTPVSNLGNIMFAPTLWARMILMTHFFSVLIARDQLEQALDEIYGCAY  
 SIEPLDLPPIQL"

## ORIGIN

Alignment Scores:			
Pred. No.:	1.26e-199	Length:	7310
Score:	3574.00	Matches:	672
Percent Similarity:	99.42%	Conservative:	10
Best Local Similarity:	97.96%	Mismatches:	4
Query Match:	98.78%	Indels:	0
DB:	14	Gaps:	0
US-09-930-591-2 (1-686) x HPCPOLYP (1-7310)			
Qy	1	MetAlaProIleThrAlaTyrAlaGlnGlnThrArgGlyLeuLeuGlyCysIleIleThr	20
Db	1728	CTGGCGCCCATCAGCGGTACGCCACAGACAGGGCGCTCTTAGGGTGATATATCACC	1787
Qy	21	SerLeuThrGlyArgAspLysAsnGlnValGluGlyGluValGlnIleValSerThrAla	40
Db	1788	AGCCTAACTGCGCGGACAAAAACCAAGTGAGGGTCCAGATGTGTCTCAACTGCT	1847
Qy	41	AlaGlnThrPheLeuAlaThrCysIleAsnGlyValCysTrpThrValThyHisGlyAla	60
Db	1848	GCCCAAACTCTCTGGCAACGTGCATCAATGGGGTGTGCTGAGCTGTCTACACGGGGCC	1907
Qy	61	GlyThrArgThrIleAlaSerProLysGlyProValIleGlnMetTyrThrAsnValAsp	80
Db	1908	GGAACGAGGACCATCGCGTCACCCAGGGTCTCTGTCATCCAGATGTATACCAATGTAGAC	1967
Qy	81	GlnAspLeuValGlyTrpProAlaProGlnGlyAlaArgSerLeuThrProCysThrCys	100

```
Db 1968 CAAGACCTTGTGGGCTGGCCGCTCCCGAAGGTAGCCGCTCAITTGACACCCCTGCACCTGC 2027
Qy 101 GlySerSerAspLeuTyrLeuValThrArgHisAlaAspValIleProValArgArgArg 120
Db 2028 GGTCTCTCGGACCTTTTACCTTGGTTCACAGGACGCCGATGTCTATTCCTCGTCCGCGCGG 2087
Qy 121 GlyAspGlyArgGlySerLeuLeuSerProArgProIleSerTyrLeuLysGlySerSer 140
Db 2088 GGTGATAGCAGGGGACCTGCTGTGCGCCCGGCCCATTTCTTACTTGAAGGCTCTCTCG 2147
Qy 141 GlyGlyProLeuLeuCysProAlaGlyHisAlaValGlyIlePheArgAlaAlaValCys 160
Db 2148 GGGGGTCCGCTGTGTGTCGCCCGGGGGCACCGCTGGGCGCATATTTAGGGCGCGGCTGTC 2207
Qy 161 ThrArgGlyValAlaLysAlaValAspPheIleProValGluSerLeuGluThrThrMet 180
Db 2208 ACCGTGGAGTGGCTAAGGCGGTGGACTTTATCCTCTGTGGAGAACCTTAGACACAACCATG 2267
Qy 181 ArgSerProValPheSerAspAsnSerSerProProAlaValProGlnSerTyrGlnVal 200
Db 2268 AGTCCCGGTGTTCACGGATAACTCTCTCCACCATGACTGCCCCAGAGCTTCCAGGTG 2327
Qy 201 AlaHisLeuHisAlaProThrGlySerGlyLysSerThrLysValProAlaAlaTyrAla 220
Db 2328 GCTCACCTCCATGCTCCACAGGCAGCGGCAAAAGCACCAAGGTCCCGGCTGCATATGCA 2387
Qy 221 AlaGlnGlyTyrLysValLeuValLeuAsnProSerValAlaAlaThrMetGlyPheGly 240
Db 2388 GCTCAGGGCTATTAAGGTGCTAGTACTCAACCCCTCTGTTGCTGCAACACTGGGGCTTTGGT 2447
Qy 241 AlaTyrMetSerLysAlaHisGlyIleAspProAsnIleArgThrGlyValArgThrIle 260
Db 2448 GCTTACATGTCGAAGGCTCATGGATCGATCCTAACATCAGGACCGGGGTGAGACAAIT 2507
Qy 261 ThrThrGlySerProIleThrTyrSerThrTyrGlyLysPheLeuAlaAspGlyGlyCys 280
Db 2508 ACCACTGGCAGCCCATCAGTACTCCACCTACGCAAGTTCCTTGGCGACGGCGGTGC 2567
Qy 281 SerGlyGlyAlaTyrAspIleIleCysAspGluCysHisSerThrAspAlaThrSer 300
Db 2568 TCGGGGGGCGCTTATGACATTAATATTTGTGACAGTGGCCACTCCACGGATGCCAATCC 2627
Qy 301 IleLeuGlyIleGlyThrValLeuAspGlnAlaGluThrAlaGlyAlaArgLeuThrVal 320
Db 2628 ATCTTGGGCATCGCATGCTCTTGACCAAGCAGAGACTCGCGGGCGAGACTGGTTGTG 2687
Qy 321 LeuAlaThrAlaThrProProGlySerValThrValProHisProAsnIleGluGluVal 340
Db 2688 CTGCGCACCGCCACCTCCGCGGCTCCGTCACTGTGCCCCATCCCAACATCGAGGAGTT 2747
Qy 341 AlaLeuSerThrThrGlyGluIleProPheTyrGlyLysAlaIleProLeuGluAlaIle 360
Db 2748 GCTCTGTCCACCCGAGAGATCCCTTTTACGGCAGGCTATCCCTCGAAGTAATC 2807
Qy 361 LysGlyGlyArgHisLeuIlePheCysHisSerLysLysLysCysAspGluLeuAlaAla 380
Db 2808 AAGGGGGGAGACATCTCATCTTCTGTCATTCAAGAAGAGTGCAGCAACTCGCGCA 2867
Qy 381 LysLeuValAlaLeuGlyValAsnAlaValAlaTyrTyrArgGlyLeuAspValSerVal 400
Db 2868 AAGCTGTGTCATTTGGGCATCAATCGGTGGCCCTACTACCGCGGTCTTGACGTGTCGTC 2927
Qy 401 IleProThrSerGlyAspValValValAlaThrAspAlaLeuMetThrGlyPheThr 420
Db 2928 ATCCGACCCAGCGCGGATGTTGTCGTGTCGGCAACCGATGCCCTCATGCCGCTATACC 2987
Qy 421 GlyAspPheAspSerValIleAspCysAsnThrCysValThrGlnThrValAspPheSer 440
Db 2988 GGGCACTTCGACTCGGTGATAGACTGCAATACGTGTGTACCAGACAGTCCGATTTTCAGC 3047
Qy 441 LeuAspProThrPheThrIleGluThrIleThrLeuProGlnAspAlaValSerArgThr 460
Db 3048 CTTGACCCCTACCTTACCATTTGAGACAATCACGCTCCCGCAGGATGCTGTCTCCCGACT 3107
```

```
Qy 461 GlnArgArgGlyArgThrGlyArgGlyLysProGlyLysTyrArgPheValAlaProGly 480
Db 3108 CAACGTGGGGGAGGACTGGCAGGGGAAGCCAGGCACTTACAGATTTGTGGCACCGGGG 3167
Qy 481 GluArgProSerGlyMetPheAspSerSerValLeuCysGluCysTyrAspAlaGlyCys 500
Db 3168 GAGCGCCCTCCGGCATGTTGACTCGTCCGTCTCTGTGAGTGTCTATGACGAGGCTGT 3227
Qy 501 AlaTyrTyrGluLeuThrProAlaGluThrThrValArgLeuArgAlaTyrMetAsnThr 520
Db 3228 GCTTGTGTATGAGTCTACGCGCGCGAGACTTACAGTTAGGCTACGAGCTACATGAACACC 3287
Qy 521 ProGlyLeuProValCysGlnAspHisLeuGluPheTyrGluGlyValPheThrGlyLeu 540
Db 3288 CCGGGGCTTCCCGGTGGCCAGGACCATCTTGAATTTTGGAGGGCGCTTTTACAGGCGCTC 3347
Qy 541 ThrHisIleAspAlaHisPheLeuSerGlnThrLysGlnSerGlyGluAsnLeuProTyr 560
Db 3348 ACTCATATAGATGCCCATTTCTATCCAGACAAAGCAGAGTGGGGAGAACTTCCTTAC 3407
Qy 561 LeuValAlaTyrGlnAlaThrValCysAlaArgAlaGlnAlaProProSerTyrAsp 580
Db 3408 CTGGTAGCGTACCAAGCCACCGGTGTGCTAGGCGCTCAAGCCCTCCCCATCGTGGGAC 3467
Qy 581 GlnMetTyrLysCysLeuIleArgLeuLysProThrLeuHisGlyProThrProLeuLeu 600
Db 3468 CAGATGTGAAGTGTGTGATTGCTTCAAGCCACCCCTCATGGGGCAACACCCCTGCTA 3527
Qy 601 TyrArgLeuGlyAlaValGlnAsnGluValThrLeuThrHisProValThrTyrIle 620
Db 3528 TACAGACTGGGCGCTGTTTCAAGATGAATCACCCTGACGACCCAGCTCACCAAATACATC 3587
Qy 621 MetThrCysMetSerAlaAspLeuGluValValThrSerThrTyrValLeuValGly 640
Db 3588 ATGACATGATGTGCGCGCACCTGGAGGTGCTCACAGACCCCTGGGTGCTGTGGCGGC 3647
Qy 641 ValLeuAlaAlaLeuAlaAlaTyrCysLeuSerThrGlyCysValValIleValGlyArg 660
Db 3648 GTCTCGCTGCTTTGGCCGCGTATTGCTGTCAACAGGCTGGGTGCTATAGTGGGCGAG 3707
Qy 661 IleValLeuSerGlyLysProAlaIleIleProAspArgGluValLeuTyrArgGluPhe 680
Db 3708 GTGCTGTTGTCCGGGAAGCGGCAATCATACCTGACAGGGAAGTCTCTTACCGAGAGTTC 3767
Qy 681 AspGluMetGluGluCys 686
Db 3768 GATGAGATGGAAGAGTGC 3785

RESULT 13
LOCUS 108294 9185 bp DNA linear PAT 02-DEC-1994
DEFINITION Sequence 1 from Patent EP 0388232.
ACCESSION 108294
VERSION 108294.1 GI:588994
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS Houghton, M., Choo, Q.-L. and Kuo, G.
TITLE NANBV diagnostics and vaccines
JOURNAL Patent: EP 0388232-A1 1 19-SEP-1990;
FEATURES
source
1. .9185
/organism="unknown"
/mol_type="unassigned DNA"

ORIGIN

Alignment Scores:
Pred. No.: 1,67e-199 Length: 9185
Score: 3574.00 Matches: 672
Percent Similarity: 99.42% Conservative: 10
```

Best Local Similarity: 97.96%		Mismatches: 4	
Query Match: 98.78%		Indels: 0	
DB: 6		Gaps: 0	
US-09-930-591-2 (1-686) x I08294 (1-9185)			
Qy	1	MetAlaProIleThrAlaTyrAlaGlnGlnThrArgGlyLeuLeuGlyCysIleIleThr	20
Db	3395	CTGGCGCCATCACGGCGTACGCCAGCAGACAAGGGCCCTCTAGGGTGATATCAC	3454
Qy	21	SerLeuThrGlyArgAspLysAsnGlnValGluGlyGluValGlnIleValSerThrAla	40
Db	3455	AGCCTAACTGGCGGAGCAAAAACCAAGTGGAGGGTGAGGTCCAGATTTGTCAACTGCT	3514
Qy	41	AlaGlnThrPheLeuAlaThrCysIleAsnGlyValCysTrpThrValTyrHisGlyAla	60
Db	3515	GCCCAAACTTCTCTGCAACGTGCATCAATGGGGTGTGCTGACATGTCTACCAACGGGGCC	3574
Qy	61	GlyThrArgThrIleAlaSerProLysGlyProValIleGlnMetTyrThrAsnValAsp	80
Db	3575	GGAAACGAGGACCATCGCTCAACCAAGGGTCTGTCTATCCAGATGTATACCAATGTAGAC	3634
Qy	81	GlnAspLeuValGlyTrpProAlaProGlnGlyAlaArgSerLeuThrProCysThrCys	100
Db	3635	CAAGACCTTGTGGGCTGGCGCGCTCCGCAAGGTAGCCGCTCATTTGACACCCCTGCACATTGC	3694
Qy	101	GlySerSerAspLeuTyrLeuValThrArgHisAlaAspValIleProValArgArg	120
Db	3695	GGCTCTCGGACCTTACCTGCTCAGAGGACACCGCGATGTCAATTCGGTGGCGCGGGGG	3754
Qy	121	GlyAspGlyArgGlySerLeuLeuSerProArgProIleSerTyrLeuLysGlySerSer	140
Db	3755	GGTGATAGAGGGGAGCTGCTGTGCGCCCGGCCCATTTCTACTTTGAAAGGCTCCTCTCG	3814
Qy	141	GlyGlyProLeuLeuCysProAlaGlyHisAlaValGlyIlePheArgAlaAlaValCys	160
Db	3815	GGGGGTCCGCTGTGTGCCCCGGGGGACGCCGTGGGCATATTTAGGGCGCGGTGTC	3874
Qy	161	ThrArgGlyValAlaLysAlaValAspPheIleProValGluSerLeuGluThrThrMet	180
Db	3875	ACCGTGGAGTGGCTAAGCGGGTGGACTTTATCTCTGTGGAGAACCTTAGACACAACCATG	3934
Qy	181	ArgSerProValPheSerAspAsnSerSerProProAlaValProGlnSerTyrGlnVal	200
Db	3935	AGGTCCCGGTGTTCACGGAATACTCTCTCCACCAAGTAGTGCCTCCAGAGCTTCCAGGTG	3994
Qy	201	AlaHisLeuHisAlaProThrGlySerGlyLysSerThrLysValProAlaAlaTyrAla	220
Db	3995	GCTCACCCTCCATGCTCCACAGCAGCGGCAAAAAGCAACCAAGGTCCCGGCTGCATATGCA	4054
Qy	221	AlaGlnGlyTyrLysValLeuValLeuAsnProSerValAlaAlaThrMetGlyPheGly	240
Db	4055	GCTCAGGGCTATAGGTGCTAGTACTCAACCCCTCTGTGTCTGCAACACTGGGCTTTGGT	4114
Qy	241	AlaTyrMetSerLysAlaHisGlyIleAspProAsnIleArgThrGlyValArgThrIle	260
Db	4115	GCTTACATGTCCAAGGCTCATGGATCGATCTTAACATCAGGACCGGGGTGAGAACAAAT	4174
Qy	261	ThrThrGlySerProIleThrTyrSerThrTyrGlyLysPheLeuAlaAspGlyCys	280
Db	4175	ACCACTGGCAGCCCATCAGCTACTCCACCTPACGGCAAGTTCTTGGCGAGCGGGGTGC	4234
Qy	281	SerGlyGlyAlaTyrAspIleIleCysAspGluCysHisSerThrAspAlaThrSer	300
Db	4235	TCGGGGGGCGCTTATGACATAATAATTTGTGACAGGTGCCACTCCAGGATGCCACATCC	4294
Qy	301	IleLeuGlyIleGlyThrValLeuAspGlnAlaGluThrAlaGlyAlaArgLeuThrVal	320
Db	4295	ATCTTGGGCATCGGCACTGTCTTGACCAACAGCAGAGACTGGCGGGCGGAGACTGGTTGTG	4354
Qy	321	LeuAlaThrAlaThrProGlySerValThrValProHisProAsnIleGluGluVal	340
Db	4355	CTGGCCACCGCACCCCTCCGGGCTCCGTCACCTGTGCCCATCCCAACATCGAGAGGTT	4414
Qy	341	AlaLeuSerThrThrGlyGluIleProPheTyrGlyLysAlaIleProLeuGluAlaIle	360
Db	4415	GCTCTGTCCACCGGAGATCCCTTTTACGGCAAGGCTATCCCGCTCGAAGTAATC	4474
Qy	361	LysGlyGlyArgHisIleLeuIlePheCysHisSerLysLysCysAspGluLeuAlaAla	380
Db	4475	AAGGGGGGAGACATCTCATCTTCTGTCTATTCAAAAGAAAGTCGACGAACCTCGCCGCA	4534
Qy	381	LysLeuValAlaLeuGlyValAsnAlaValAlaTyrTyrArgGlyLeuAspValSerVal	400
Db	4535	AAGCTGTGCGCATTTGGGCATCAATCGCTGGCCCTACTACCGGGCTTTGACGTGCCGTC	4594
Qy	401	IleProThrSerGlyAspValValValAlaThrAspAlaLeuMetThrGlyPheThr	420
Db	4595	ATCCCGACCAACGGCGATGTGTCTGTGGCAACCGATGCCCTCATGACCGGTATACC	4654
Qy	421	GlyAspPheAspSerValIleAspCysAsnThrCysValThrGlnThrValAspPheSer	440
Db	4655	GGCGACTTCGACTCGGTGTAGACTGCAATACGTGTGTCTACCCAGACAGTCGATTTTCAGC	4714
Qy	441	LeuAspProThrPheThrIleGluThrIleThrLeuProGlnAspAlaValSerArgThr	460
Db	4715	CTTGACCTTACCTTACCATTTGACAAATCACGCTCCCCAGGATGTGTCTCCCGACT	4774
Qy	461	GlnArgArgGlyArgThrGlyArgGlyLysProGlyIleTyrArgPheValAlaProGly	480
Db	4775	CAACGTGGGGGAGGACTGGCAGGGGAGCCAGGCATCTACAGATTTGTGGCACCGGGG	4834
Qy	481	GluArgProSerGlyMetPheAspSerSerValLeuCysGluCysTyrAspAlaGlyCys	500
Db	4835	GAGCGCCCTCCGCGATGTTCGACTCGCTCGCTCCTCTGTGAGTGTCTATGACGAGGCTGT	4894
Qy	501	AlaTrpTyrGluLeuThrProAlaGluThrThrValArgLeuArgAlaTyrMetAsnThr	520
Db	4895	GCTTGGTATGAGCTCACGCCCGCGAGACTACAGTTAGGCTACAGGCTACATGAACACC	4954
Qy	521	ProGlyLeuProValCysGlnAspHisLeuGluPheTrpGluGlyValPheThrGlyLeu	540
Db	4955	CCGGGGCTTCCGCTGTGCCAGGACCATCTTGAAATTTTGGGAGGGCGTCTTTACAGGCTC	5014
Qy	541	ThrHisIleAspAlaHisPheLeuSerGlnThrLysGlnSerGlyGluAsnLeuProTyr	560
Db	5015	ACTCATATAGATGCCACTTTCTATCCAGACAAGCAGAGTGGGGAGAACCTTCTCTTAC	5074
Qy	561	LeuValAlaTyrGlnAlaThrValCysAlaArgAlaGlnAlaProProProSerTrpAsp	580
Db	5075	CTGCTAGCGTACCAGCACCGGTGCGCTAGGGCTCAAGCCCTCCCCCATCTGGGAC	5134
Qy	581	GlnMetTrpLysCysLeuIleArgLeuLysProThrLeuHisGlyProThrProLeuLeu	600
Db	5135	CAGATGTGGAAGTGTGTGATTGCTTCAAGCCCAACCTCTCCATGGCCCAACACCTTGCTA	5194
Qy	601	TyrArgLeuGlyAlaValGlnAsnGluValThrLeuThrHisProValThrLysTyrIle	620
Db	5195	TACAGACTGGCGGTGTTCAAGATGAATCACTGACGCAACCCAGTCACCAATAATCATC	5254
Qy	621	MetThrCysMetSerAlaAspLeuValValThrSerThrTrpValLeuValGlyGly	640
Db	5255	ATGACATGCATGTGGCGGACCTGGAGGTGTCACAGACACCTGGGTGCTGTGGCGGC	5314
Qy	641	ValLeuAlaAlaLeuAlaAlaTyrCysLeuSerThrGlyCysValValIleValGlyArg	660
Db	5315	GTCTGTGCTTGTGGCGCGTATTGCTGTCAACAGGCTGGGTGTCATAGTGGGAGG	5374
Qy	661	IleValLeuSerGlyLysProAlaIleIleProAspArgGluValLeuTyrArgGluPhe	680
Db	5375	GTGCTCTTGTCCGGGAAGCGGCAATCATACCTGACAGGGAAGTCTCTTACCGAGAGTTC	5434
Qy	681	AspGluMetGluGluCys	686
Db	5435	GATGAGATGGAAGAGTGC	5452

```

RESULT 14
LOCUS      BD091382                9185 bp    DNA        linear    PAT 27-AUG-2002
DEFINITION HCV cultivation method in eucaryotic cells.
ACCESSION  BD091382
VERSION    BD091382.1 GI:22636993
KEYWORDS   JP 2001314192-A/3.
SOURCE     Homo sapiens (human)
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE  1 (bases 1 to 9185)
AUTHORS   Weiner,A.J., Steimer,K.S. and Houghton,M.
TITLE     HCV cultivation method in eucaryotic cells
JOURNAL   Patent: JP 2001314192-A '3 13-NOV-2001;
          CHIRON CORP
COMMENT   OS Homo sapiens (human)
          PN JP 2001314192-A/3
          PD 13-NOV-2001
          PF 15-MAR-2001 JP 2001075114
          PR 25-AUG-1989 US 398667
          PI AMY J WEINER,KATHELYN S STEIMER,MICHAEL HOUGHTON PC
          CI2N15/09,C12N5/10,C12N7/00//C12N1/00,C12R1/93,C12N15/00, PC
          CI2N5/00
          CC HCV cultivation method in eucaryotic cells
          FH Key
          FT source
          FT Location/Qualifiers
          FT 1. 9185
          FT /organism='Homo sapiens (human)'.

FEATURES
     source              Location/Qualifiers
     1. 9185
        /organism="Homo sapiens"
        /mol_type="genomic DNA"
        /db_xref="taxon:9606"

ORIGIN
Alignment Scores:
Pred. No.:      1.67e-199      Length:      9185
Score:          3574.00        Matches:    672
Percent Similarity: 99.42%      Conservative: 10
Best Local Similarity: 97.96%    Mismatches:  4
Query Match:     98.78%         Indels:     0
DB:              6              Gaps:       0

US-09-930-591-2 (1-686) x BD091382 (1-9185)

QY      1 MetAlaProIleThrAlaTyrAlaGlnGlnThrArgGlyLeuLeuGlyCysIleIleThr 20
Db      ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
3395   CTGGCGCCCATCAGCGGTACGCCCAAGCAAGCGGCGCTCTTAGGGTGCATAATCACC 3454

QY      21 SerLeuThrGlyArgAspLysAsnGlnValGluGlyValGlnIleValSerThrAla 40
Db      ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
3455   AGCTTAATGCGCGGACAAAACCAAGTGGAGGGTGAGTCCAGATTGTGTCACTGCT 3514

QY      41 AlaGlnThrPheLeuAlaThrCysIleAsnGlyValCysTrpThrValTyrHisGlyAla 60
Db      ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
3515   GCCCAAACTTCTCTGGCAAGTGCATCAATGGGGTGTCTGGACTGTCTACACGGGGCC 3574

QY      61 GlyThrArgThrIleAlaSerProLysGlyProValIleGlnMetTyrThrAsnValAsp 80
Db      ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
3575   GGAACGAGGACCATCGCGTCACCAAGGGTCTGTATCCAGATGATACCAATGATAGAC 3634

QY      81 GlnAspLeuValGlyTrpProAlaProGlnGlyAlaArgSerLeuThrProCysThrCys 100
Db      ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
3635   CAAGACCTTGTGGGTGGCCCGCTCGCAAGGTAGCGGTCAATGACACCTGCATTTGC 3694

QY      101 GlySerSerAspLeuTyrLeuValThrArgHisAlaAspValIleProValArgArg 120
Db      ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
3695   GGCTCTCTGGACCTTTACTTGGTTCAGGACGCGCATGTCTATCCCGTGGCGCGCG 3754

QY      121 GlyAspGlyArgGlySerLeuLeuSerProArgProIleSerTyrLeuLysGlySerSer 140
Db      ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
3755   GGTGATAGCAGGGGCGCCTGCTGTGCCCCCGCCCATTTCTTCTTACATTGAAGGCTCTCTCG 3814

```

```

141  GlyGlyProLeuLeuCysProAlaGlyHisAlaValGlyIlePheArgAlaAlaValCys 160
161  ThrArgGlyValAlaLysAlaValAspPheIleProValGluSerLeuGluThrThrMet 180
3875  ACCGTGGAGTGGCTAAGCGGTGGACTTTATCCCTGTGTGAGAAACCTAGAGAACCAATG 3934
181  ArgSerProValPheSerAspAsnSerSerProAlaValProGlnSerTyrGlnVal 200
3935  AGTCCCCCGGTGTTCACGGATAACTCTCTCCACCAAGTAGTCCCGCAGAGCTTCCAGGTG 3994
201  AlaHisLeuHisAlaProThrGlySerGlyLysSerThrLysValProAlaAlaTyrAla 220
3995  GCTCACTCCATGCTCCACAGCAGCGGCAAAAGCACCAGGTCCCGGTGCATATGCA 4054
221  AlaGlnGlyTyrLysValLeuValLeuAsnProSerValAlaAlaThrMetGlyPheGly 240
4055  GCTCAGGGCTATAAAGGTGCTAGTACTCAACCCCTCTGTGTGTCACACACTGGGCTTTGGT 4114
241  AlaTyrMetSerLysAlaHisGlyIleAspProAsnIleArgThrGlyValArgThrIle 260
4115  GCTTACATGTCGAAGGCTCATGGATCGATCTTAACATCAGGACCGGGGTGAGAAACAAT 4174
261  ThrThrGlySerProIleThrTyrSerThrTyrGlyLysPheLeuAlaAspGlyGlyCys 280
4175  ACCACTGGCAGCCCATCATCAGTACTCCACTACGGCAAGTTCCTTGCCGACGGCGGTGC 4234
281  SerGlyGlyAlaTyrAspIleIleLeuLeuCysAspGluCysHisSerThrAspAlaThrSer 300
4235  TCGGGGGCGGCTTATGACATAATAATTTGTGACGAGTGCCTCCACGGATGCAATCC 4294
301  IleLeuGlyIleGlyThrValLeuAspGlnAlaGluThrAlaGlyAlaArgLeuThrVal 320
4295  ATCTTGGGCATCGGCACATGCTCTTGACCAAGCAGAGACTGCGGGGCGAGACTGGTTGTG 4354
321  LeuAlaThrAlaThrProGlySerValThrValProHisProAsnIleGluGluVal 340
4355  CTCGCCACCGCCACCCCTCCGGGCTCCGTCACCTGTGCCCCCATCCCAACATCAGGAGGT 4414
341  AlaLeuSerThrThrGlyGluIleProPheTyrGlyLysAlaIleProLeuGluAlaIle 360
4415  GCTCTGTCCACCACCGGAGAGATCCCTTTTACGGCAAGGCTATCCCTCCGAAGTAATC 4474
361  LysGlyGlyArgHisLeuIlePheCysHisSerLysLysLysCysAspGluLeuAlaAla 380
4475  AAGGGGGGAGACATCTCATCTTGTCTATTCAAGAAGAGTGCAGCAATCTCGCGCA 4534
381  LysLeuValAlaLeuGlyValAsnAlaValAlaTyrTyrArgGlyLeuAspValSerVal 400
4535  AAGTGGTGGCATTTGGGCATCAATGCGGTGGCTACTACCGCGTCTTACGTGTCGTC 4594
401  IleProThrSerGlyAspValValValAlaThrAspAlaLeuMetThrGlyPheThr 420
4595  ATCCCGACACGAGCGGATGTTGTCTGCTGGCAACCGCATCGCCCTCATGACCGGTATACC 4654
421  GlyAspPheAspSerValIleAspCysAsnThrCysValThrGlnThrValAspPheSer 440
4655  GCGCACTTCGATCGGTATAGACTGCATATCGTGTGTCCACCAAGACAGTGCATTTCAAC 4714
441  LeuAspProThrPheThrIleGluThrIleThrLeuProGlnAspAlaValSerArgThr 460
4715  CTTGACCTTACCTTCAACATTGAGACAATCACGCTCCCCCAGGATGCTGTCTCCGCACT 4774
461  GlnArgArgGlyArgThrGlyArgGlyLysProGlyIleTyrArgPheValAlaProGly 480
4775  CAACGTTCGGGGCAGGACTGCGCAGGGGAGCGCAGGCACTACAGATTTTGTGCACCGGG 4834
481  GluArgProSerGlyMetPheAspSerValLeuCysGluCysTyrAspAlaGlyCys 500
4835  GAGCGCCCCCTCGGATATGTTTCAGTCTGCTCCCTCTGTGAGTGTATGACGAGGCTGT 4894
501  AlaTrpTyrGluLeuThrProAlaGluThrThrValArgLeuArgAlaTyrMetAsnThr 520

```

```
Db 4895 GCTTGGTATAGCTCACCCCGGAGACTACAGTTAGGCTACGAGCGTACATGAACACC 4954
Qy 521 ProGlyLeuProValCysGlnAspHisLeuGluPheTrpGluGlyValPheThrGlyLeu 540
Db 4955 CCGGGCTTCCTGGTGGCAGGACCATCTGCAATTTGGGAGGGCGTCTTTACAGGCTC 5014
Qy 541 ThrHisLeuAspAlaHisPheLeuSerGlnThrLysGlnSerGlyGluAsnLeuProTyr 560
Db 5015 ACTCATATAGATGCCCACTTTCTATCCAGACAAAGCAGAGTGGGGAGAACCTTCTTAC 5074
Qy 561 LeuValAlaTyrGlnAlaThrValCysAlaArgAlaGlnAlaProProSerTrpAsp 580
Db 5075 CTGTAGCGTACCAAGCACCCTGTCCTAGGCGCTCAAGCCCTCCCCCATCGTGGAC 5134
Qy 581 GlnMetTrpLysCysLeuIleArgLeuLysProThrLeuHisGlyProThrProLeuLeu 600
Db 5135 CAGATGTGGAAGTGTGTTGATTCGCTCAAGCCACCTCCATGGGCCACACCTCTGCTA 5194
Qy 601 TyrArgLeuGlyAlaValGlnAsnGluValThrLeuThrHisProValThrLysTyrIle 620
Db 5195 TACAGCTGGCGCTGTTCAGAAATCAACCTGACGCCACCCAGTCACCAATAATCATC 5254
Qy 621 MetThrCysMetSerAlaAspLeuValValThrSerThrTrpValLeuValGlyGly 640
Db 5255 ATGACATGCATGTCCGGCGACCTGGAGGTGTCACGAGCACCTGGGTGCTCGTTGGCGGC 5314
Qy 641 ValLeuAlaAlaLeuAlaAlaTyrCysLeuSerThrGlyCysValValIleValGlyArg 660
Db 5315 GTCTGGCTGTTTGGCGCGTATGTGCTGCAAGGCTCGTGGTCATAGTGGCAGG 5374
Qy 661 IleValLeuSerGlyLysProAlaIleIleProAspArgGluValLeuTyrArgGluPhe 680
Db 5375 GTCTGCTGTCCGGAGCGGCAATCATACCTGACAGGGAAGTCTCTACCGAGAGTTC 5434
Qy 681 AspGluMetGluGluCys 686
Db 5435 GATGAGATGGAAGAGTGC 5452

RESULT 15
LOCUS AR166930 9379 bp DNA linear PAT 17-OCT-2001
DEFINITION Sequence 1 from patent US 6284249.
ACCESSION AR166930
VERSION AR166930.1 GI:16243325
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 9379)
AUTHORS Barban,V.
TITLE Fusion polypeptide having the C protein and E1 protein of hepatitis
C virus
JOURNAL Patent: US 6284249-A 1 04-SEP-2001;
FEATURES
    Location/Qualifiers
        1..9379
            /organism="unknown"
            /mol_type="unassigned DNA"

ORIGIN

Alignment Scores:
Pred. No.: 1.71e-199 Length: 9379
Score: 3574.00 Matches: 672
Percent Similarity: 99.42% Conservative: 10
Best Local Similarity: 97.96% Mismatches: 4
Query Match: 98.78% Indels: 0
DB: 6 Gaps: 0

US-09-930-591-2 (1-686) x AR166930 (1-9379)

Qy 1 MetAlaProIleThrAlaTyrAlaGlnGlnThrArgGlyLeuLeuGlyCysIleIleThr 20
Db 3395 CTGGCGGCCCATCACGGCGGTACGCCGACAGACAAAGGGGCGCTCTTAGGGTGCATAATCACC 3454
```

```
Qy 21 SerLeuThrGlyArgAspLysAsnGlnValGluGlyGluValGlnIleValSerThrAla 40
Db 3455 AGCCTAACTGCCGGGCAAAAACCAAGTGAGGGTGAGGTCCAGATTGTGTCAACTGCT 3514
Qy 41 AlaGlnThrPheLeuAlaThrCysIleAsnGlyValCysTrpThrValTyrHisGlyAla 60
Db 3515 GCCCAAACTTCTCGGCAACGTCATCAATGGGGTGTCTGAGACTGTCTACACGGGGCC 3574
Qy 61 GlyThrArgThrIleAlaSerProLysGlyProValIleGlnMetTyrThrAsnValAsp 80
Db 3575 GGAACGAGGACCATCGCGTCAACCAAGGTCCTGTCTATCCAGATGTATACCAATGTAGAC 3634
Qy 81 GlnAspLeuValGlyTrpProAlaProGlnGlyAlaArgSerLeuThrProCysThrCys 100
Db 3635 CAAGACCTTGTGGGCTGGCCGCTCCGCAAGTAGCGCTCATTGACACCCCTGCACATTGC 3694
Qy 101 GlySerSerAspLeuTyrLeuValThrArgHisAlaAspValIleProValArgArgArg 120
Db 3695 GGCTCTCGGACCTTTACCTGGTCACGAGCACCGCATGTCTATCCCGTGCAGCGCGG 3754
Qy 121 GlyAspGlyArgGlySerLeuLeuSerProArgProIleSerTyrLeuLysGlySerSer 140
Db 3755 GGTGATAGCAGGGGACGCTGTCTGCCCGGCCCATTTCTTACTTGAAGGCTCTCTCG 3814
Qy 141 GlyGlyProLeuLeuCysProAlaGlyHisAlaValGlyIlePheArgAlaAlaValCys 160
Db 3815 GGGGTCCTCGTGTGTGCCCGGGGACGCGTGGGCATATTTAGGGCGCGGTGTGC 3874
Qy 161 ThrArgGlyValAlaLysAlaValAspPheIleProValGluSerLeuGluThrThrMet 180
Db 3875 ACCCGTGGAGTGGCTAAGCGGTGGAATTTATCTCTGTGGAGAACCTTAGACAAACCATG 3934
Qy 181 ArgSerProValPheSerAspAsnSerSerProProAlaValProGlnSerTyrGlnVal 200
Db 3935 AGGTCCCGGTGTTCACGGAATACTCTCTCCACAGTAGTCCCGACAGCTTCCAGGTG 3994
Qy 201 AlaHisLeuHisAlaProThrGlySerGlyLysSerThrLysValProAlaAlaTyrAla 220
Db 3995 GCTCACCTCCATGCTCCACAGGCGGCAAGACCAAGGTCCCGGTGCATATGCA 4054
Qy 221 AlaGlnGlyTyrLysValLeuValLeuAsnProSerValAlaAlaThrMetGlyPheGly 240
Db 4055 GCTCAGGGCTATAAGGTGTAGTACTCAACCCCTCTGTGTGTGCAACACTGGGTGTGGT 4114
Qy 241 AlaTyrMetSerLysAlaHisGlyIleAspProAsnIleArgThrGlyValArgThrIle 260
Db 4115 GCTTACATGTCCAAGGCTCATGGATCGATCTTAACATCAGGACCGGGGTGAGAACAT 4174
Qy 261 ThrThrGlySerProIleThrTyrSerThrTyrGlyLysPheLeuAlaAspGlyGlyCys 280
Db 4175 ACCACTGGCAGCCCATCAGTACTCCACCTACGGCAAGTTCCTTGGCGAGCGGGTGC 4234
Qy 281 SerGlyGlyAlaTyrAspIleIleCysAspGluCysHisSerThrAspAlaThrSer 300
Db 4235 TCGGGGGCGCTTATGACATAATAATTTGTACAGAGTGCACATCCACGGATGCCACATCC 4294
Qy 301 IleLeuGlyIleGlyThrValLeuAspGlnAlaGluThrAlaGlyAlaArgLeuThrVal 320
Db 4295 ATCTTGGGCATCGGCACTGTCTTGACCAAGCAGAGACTCGGGGGGCGAGACTGGTGTG 4354
Qy 321 LeuAlaThrAlaThrProProGlySerValThrValProHisProAsnIleGluGluVal 340
Db 4355 CTGGCACCGCCACCTCCGGGCTCGTCACTGTGCCCCATCCCAACATCGAGGAGTT 4414
Qy 341 AlaLeuSerThrThrGlyGluIleProPheTyrGlyLysAlaIleProLeuGluAlaIle 360
Db 4415 GCTCTGTCCACACCGGAGAGATCCCTTTTACGGCAAGGCTATCCCCCTCGAAGTAATC 4474
Qy 361 LysGlyGlyArgHisLeuIlePheCysHisSerLysLysCysAspGluLeuAlaAla 380
Db 4475 AAGGGGGGAGAGATCTCATCTTCTGTCTTCAAGAAAGAGTGCAGCAACTCGCGCA 4534
```

Qy	381	LysLeuValAlaLeuGlyValAsnAlaValAlaTyrTyrArgGlyLeuAspValSerVal	400
Db	4535	AAAGCTGGTCGCATTGGGCATCAATGCGGTGGCTACTACCGGGTCTTGACGTGTCGGTC	4594
Qy	401	IleProThrSerGlyAspValValValAlaThrAspAlaLeuMetThrGlyPheThr	420
Db	4595	ATCCCGACGAGCGCGCATGTTGTCGTGGGACCGGATGCCCTCATGACCGGCTATACC	4654
Qy	421	GlyAspPheAspSerValIleAspCysAsnThrCysValThrGlnThrValAspPheSer	440
Db	4655	GGCGACTTCGACTCGGTGATAGACTGCAATACGTGTGTCCACGACAGTCGATTTTCAGC	4714
Qy	441	LeuAspProThrPheThrIleGluThrIleThrLeuProGlnAspAlaValSerArgThr	460
Db	4715	CTTGACCTTACCTTCACCATTTAGACAATCACGCTCCCCAGGATGCTGTCTCCCGCACT	4774
Qy	461	GlnArgGlyArgThrGlyArgGlyLysProGlyIleTyrArgPheValAlaProGly	480
Db	4775	CAACGTGGGGCAGGACTGGCAGGGGGAGCCAGGCATCTACAGATTTGTGGCACCAGGG	4834
Qy	481	GluArgProSerGlyMetPheAspSerSerValLeuCysGluCysTyrAspAlaGlyCys	500
Db	4835	GAGCGCCCTCCGGCATGTTTCGACTCGTCGTCCTCTGTGAGTGTATGACGAGGCTGT	4894
Qy	501	AlaTrpTyrGluLeuThrProAlaGluThrThrValArgLeuArgAlaTyrMetAsnThr	520
Db	4895	GCTTGGTATGAGTTCACGCCCGCGAGACTACAGTTAGGCTACGAGGTACATGAACACC	4954
Qy	521	ProGlyLeuProValCysGlnAspHisLeuGluPheTrpGluGlyValPheThrGlyLeu	540
Db	4955	CCGGGGCTTCCCGTGTGCCAGGACCATCTTGAATTTGGGAGGGCGTCTTTACAGGCCTC	5014
Qy	541	ThrHisIleAspAlaHisPheLeuSerGlnThrLysGlnSerGlyGluAsnLeuProTyr	560
Db	5015	ACTCATATAGATGCCACTTTCTATCCAGACAAAGCAGAGTGGGGAGAAACCTTCCTTAC	5074
Qy	561	LeuValAlaTyrGlnAlaThrValCysAlaArgAlaGlnAlaProProSerTrpAsp	580
Db	5075	CTGGTAGCGTACCAAGCACCGTGTGCTTAGGGCTCAAGCCCTCCCCCATCGTGGGAC	5134
Qy	581	GlnMetTrpLysCysLeuIleArgLeuLysProThrLeuHisGlyProThrProLeuLeu	600
Db	5135	CAGATGTGGAAGTGTGTTGATTGGCTCAAGCCCAACCCCTCATGGGCCAACACCCCTGCTA	5194
Qy	601	TyrArgLeuGlyAlaValGlnAsnGluValThrLeuThrHisProValThrLysTyrIle	620
Db	5195	TACAGACTGGGGCGCTGTTCAGAAATGAAATCACCTGACGACCCCGCCAGTCACCAATACATC	5254
Qy	621	MetThrCysMetSerAlaAspLeuGluValValThrSerThrTrpValLeuValGlyGly	640
Db	5255	ATGACATGCATGTCGGCCGACCTGGAGGTGTCACAGGACCTGGGTGCTCGTTGGCGGC	5314
Qy	641	ValLeuAlaAlaLeuAlaAlaTyrCysLeuSerThrGlyCysValValIleValGlyArg	660
Db	5315	GTCCTGGCTGCTTTGGCGCGCTATTGCTGTCAACAGGCTGCGTGGTCATAGTGGGAGG	5374
Qy	661	IleValLeuSerGlyLysProAlaIleLeuProAspArgGluValLeuTyrArgGluPhe	680
Db	5375	GTCGCTTGTCCGGGAGCGGCAATCATACCTGACAGGGAAGTCCCTCTACCGAGAGTTTC	5434
Qy	681	AspGluMetGluGluCys	686
Db	5435	GATGAGATGGAGAGTGC	5452

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: February 27, 2005, 01:21:41 ; Search time 834 Seconds  
(without alignments)  
4869.236 Million cell updates/sec

Title: US-09-930-591-2

Perfect score: 3618

Sequence: 1 MAPTAYAQTRGLGCIIT.....PAIPDREVLRYFDEMEEC 686

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n model -DEV=xlp  
-Q=/cgn2\_1/USPTO\_spool\_p/US09930591/runat\_25022005\_143612\_25666/app\_query.fasta\_1.839  
-DB=N\_Geneseq\_16Dec04 -QMT=fastcap -SUFFIX=ring -MINMATCH=0.1 -LOOPCL=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi  
-LIST=45 -LOCAL=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFWT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US09930591.ecgn 1 1 708 @runat\_25022005\_143612\_25666 -NCPU=6 -ICPU=3  
-NO\_WMAP -LARGQUERY -NEG\_SCORES=0 -WAIT -DSPLOCK=100 -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPOPT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPOPT=0.5 -DELOP=6 -DELEXT=7

Database :

N\_Geneseq\_16Dec04:\*  
1: Geneseqn1980s:\*  
2: Geneseqn1990s:\*  
3: Geneseqn2000s:\*  
4: Geneseqn2001as:\*  
5: Geneseqn2001bs:\*  
6: Geneseqn2002as:\*  
7: Geneseqn2002bs:\*  
8: Geneseqn2003as:\*  
9: Geneseqn2003bs:\*  
10: Geneseqn2003cs:\*  
11: Geneseqn2003ds:\*  
12: Geneseqn2004as:\*  
13: Geneseqn2004bs:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3618	100.0	2061	6 AAD34500	Aad34500 Hepatitis
2	3618	100.0	2061	6 AAD31767	Aad31767 Hepatitis
3	3618	100.0	2061	10 AAD60868	Aad60868 Hepatitis
4	3618	100.0	2061	12 ADG47658	Adg47658 HCV NS3/4
5	3618	100.0	2078	12 ADG47692	Adg47692 HCV NS3/4

6	3574	98.8	7310	1	AAN92106	Combined
7	3574	98.8	7310	1	AAN90336	Composite
8	3574	98.8	7310	1	AAQ98221	Hepatitis
9	3574	98.8	8316	3	AA75296	cDNA sequ
10	3574	98.8	9133	2	AAZ07656	Nucleotid
11	3574	98.8	9185	2	AAQ10566	Hepatitis
12	3574	98.8	9185	3	AAA75297	Sense str
13	3574	98.8	9185	12	ADN35979	HCV cDNA
14	3574	98.8	9401	2	AAT12710	Hepatitis
15	3574	98.8	9401	2	AAT99981	HCV poly
16	3574	98.8	9401	2	AAV09989	HCV poly
17	3574	98.8	9401	6	AAD35043	Hepatitis
18	3574	98.8	9401	12	ADL23106	Hepatitis
19	3574	98.8	9401	13	ADR29358	Hepatitis
20	3571	98.7	6905	1	AAN92103	Combined
21	3571	98.7	9185	2	AAX26737	Nucleotid
22	3571	98.7	9185	2	ADF66068	Hepatitis
23	3570	98.7	9185	2	AAQ05956	Sense str
24	3570	98.7	9618	11	ADN33102	Hepatitis
25	3568	98.6	6299	4	AAF83669	HCV NS3/4
26	3568	98.6	9185	2	AAQ00459	Hepatitis
27	3567	98.6	8316	2	AAQ05955	Hepatitis
28	3565	98.5	2058	6	ABK15344	Hepatitis
29	3565	98.5	2058	6	AAD29795	HCV-1 NS3
30	3565	98.5	2058	10	ADC06768	HCV mutan
31	3565	98.5	2058	10	ABX14410	DNA encod
32	3565	98.5	2058	12	ADL66804	HCV NS3/4
33	3565	98.5	9400	2	AAQ21744	Compiled
34	3555	98.3	5360	1	AAN90327	Hepatitis
35	3552	98.2	5300	1	AAN92097	Hepatitis
36	3551	98.1	9646	2	AAV59361	Hepatitis
37	3551	98.1	9646	6	ABK87285	cDNA enco
38	3551	98.1	9646	8	ACA62466	HCV H77 c
39	3551	98.1	12980	2	AAV59364	Hepatitis
40	3551	98.1	12980	6	ABK87286	Hepatitis
41	3551	98.1	12980	8	ACA62469	DNA encod
42	3551	98.1	16622	3	RAZ36212	Nucleotid
43	3550	98.1	9502	2	AAQ74770	Hepatitis
44	3549	98.1	8316	12	ADN35977	HCV cDNA
45	3544	98.0	9518	5	AAD03778	Hepatitis

ALIGNMENTS

RESULT 1  
AAD34500  
ID AAD34500 standard; DNA; 2061 BP.  
XX  
AC AAD34500;  
XX  
DT 16-JUL-2002 (first entry)  
XX  
DE Hepatitis C virus NS3/4A protein encoding DNA.  
XX  
KW Hepatitis C virus; HCV; NS3/4A protein; therapy; HCV infection; vaccine;  
XX  
KW virucide; gene; ds.  
XX  
OS Hepatitis C virus.  
XX  
FH Key  
FT CDS  
FT Location/Qualifiers  
FT 1..2061  
FT /\*tag= a  
FT /product= "HCV NS3/4A protein"  
XX  
PN WO200214362-A2.  
XX  
PD 21-FEB-2002.  
XX  
PF 15-AUG-2001; 2001WO-IB001774.  
XX  
PR 17-AUG-2000; 2000US-0225767P.  
PR 29-AUG-2000; 2000US-0229175P.  
PR 03-NOV-2000; 2000US-00705547.



XX (TRIP-) TRIPEP AB.  
 XX PA Sallberg M;  
 XX PI WPI; 2002-339446/37.  
 XX DR P-PSDB; AAE21837.  
 XX PT Novel hepatitis C virus NS3/4A peptide useful for diagnosing presence or  
 XX PT absence of hepatitis C virus in a subject and for preparing a medicament  
 XX PT for treating hepatitis C virus infection.  
 XX PS Claim 1; Page 64-65; 90pp; English.  
 XX CC The present invention relates to novel hepatitis C virus (HCV) NS3/4A  
 CC proteins and their corresponding polynucleotides. NS3/4A sequences are  
 CC useful for identifying the presence or absence of HCV in a subject. They are  
 CC useful for preparing a medicament used for treating or preventing HCV  
 CC infection. Sequences of the invention are also used as vaccines. The  
 CC present sequence is a DNA encoding HCV NS3/4A protein  
 XX CC  
 XX SQ Sequence 2061 BP; 427 A; 616 C; 571 G; 447 T; 0 U; 0 Other;  
 XX

Alignment Scores:  
 Pred. No.: 3,22e-245 Length: 2061  
 Score: 3618.00 Matches: 686  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 6 Gaps: 0

US-09-930-591-2 (1-686) x AAD34500 (1-2061)

QY 1 MetAlaProIleThrAlaValAlaGlnGlnThrArgGlyLeuLeuGlyCysIleIleThr 20  
 DB 1 ATGGCGCTATACAGGCTATGCCAGACAGAGGGGCTTTGGGATGCATAATCACC 60  
 QY 21 SerLeuThrGlyArgAspLysAsnGlnValGluGlyValGlnIleValSerThrAla 40  
 DB 61 AGCTTGACCGCGCGGACAAAACACAGGTGAGGGTGAGGTTTCAGATCGTGCACTGCT 120  
 QY 41 AlaGlnThrPheLeuAlaThrCysIleAsnGlyValCysTrpThrValThrHisGlyAla 60  
 DB 121 GCCCAGACTTCTTGGCAACTGCTAATACGGGGTGTTGGAGCTGTCTACCATGGAGCC 180  
 QY 61 GlyThrArgThrIleAlaSerProLysGlyProValIleGlnMetTyrThrAsnValAsp 80  
 DB 181 GGAAACAGGACCATTTGGCTCACCTAAGGGTCCTGTTATCCAGATGTACCAATGTGGAC 240  
 QY 81 GlnAspLeuValGlyTrpProAlaProGlnGlnGlyAlaArgSerLeuThrProCysThrCys 100  
 DB 241 CNAGACCTCGTAGGCTGGCCCGCTCCCAAGTGCCCGCTCATTAACACCATGCACCTGC 300  
 QY 101 GlySerSerAspLeuThrLeuValThrArgHisAlaAspValIleProValArgArgArg 120  
 DB 301 GGCTCTCTCGGACCTTTACCTGGTGCACGAGCAGCGCGATGTCATTCTGTGCGCGCAGCG 360  
 QY 121 GlyAspGlyArgGlySerLeuLeuSerProArgProIleSerTyrLeuLysGlySerSer 140  
 DB 361 GGTGATGGCAGGGGACCGCTTTCGCCCGGCTATCTTACTTTGAAGGCTCTCTCG 420  
 QY 141 GlyGlyProLeuLeuCysProAlaGlyHisAlaValGlyIlePheArgAlaAlaValCys 160  
 DB 421 GGAGGCCCTCTGTGTGCCCGCGAGACATGCCGTAGGCATATTCAGAGCGCGGTATGC 480  
 QY 161 ThrArgGlyValAlaLysAlaValAspPheIleProValGluSerLeuGluThrThrMet 180  
 DB 481 ACCCGTGGAGTGGCTAAGCGGTGGACTTCATCCCGGTAGAGAGCTTAGAGACAACCATG 540  
 QY 181 ArgSerProValPheSerAspAsnSerSerProAlaValProGlnSerTyrGlnVal 200  
 DB 541 AGGTCCCGGTGTCTCAGACAACTCTCTCCCAACAGAGTGCCCGACAGCTACCAAGTG 600

QY 201 AlaHisLeuHisAlaProThrGlySerGlyLysSerThrLysValProAlaAlaTyrAla 220  
 DB 601 GCCCAGCTGTCATGCTCCACCGGCTAGAGCAGCAACCAAGTCCCGCGCATACGCA 660  
 QY 221 AlaGlnGlyTyrLysValLeuValLeuAsnProSerValAlaAlaThrMetGlyPheGly 240  
 DB 661 GCTCAGGGGTACAGGTGGTGGCTCAACCCCTCGTGTGCTGCAACAATGGGCTTTGGT 720  
 QY 241 AlaTyrMetSerLysAlaHisGlyIleAspProAsnIleArgThrGlyValAlaThrThrIle 260  
 DB 721 GCTTACATGTCGAAGGCCCATGGATGATCTTAACATCAGACTGGGTGAGACAAATT 780  
 QY 261 ThrThrGlySerProIleThrTyrSerThrTyrGlyLysPheLeuAlaAspGlyGlyCys 280  
 DB 781 ACTACTGGCAGCCGATCAGTATTCCACTACGCAAGTTCCTTCCGACGCGGCTGT 840  
 QY 281 SerGlyGlyAlaTyrAspIleIleCysAspGluCysHisSerThrAspAlaThrSer 300  
 DB 841 TCAGGGGGTGTCTTATGACATTAATTTGTGACAGTGCACCTCCACGGATGCAACATCC 900  
 QY 301 IleLeuGlyIleGlyThrValLeuAspGlnAlaGluThrAlaGlyAlaArgLeuThrVal 320  
 DB 901 ATCTTGGGCAATTGGCACTGCTTTCACCAAGCAGAGACCGCGGGGGGAGACTGACTGTG 960  
 QY 321 LeuAlaThrAlaThrProProGlySerValThrValProHisProAsnIleGluGluVal 340  
 DB 961 CTGCGCACCGCTACCCCTCCGGCTCCGTCACCTGTGCCCATCTTAACATCAGAGGAGTT 1020  
 QY 341 AlaLeuSerThrThrGlyGluIleProPheTyrGlyLysAlaIleProLeuGluAlaIle 360  
 DB 1021 GCTCTGTCCACTACCGGAGAGATCCCTTTTATGGCAAGGCTATCCCTTTGAAGCAATT 1080  
 QY 361 LysGlyGlyArgHisLeuIlePheCysHisSerLysLysLysCysAspGluLeuAlaAla 380  
 DB 1081 AAGGGGGGAGACATCTCATCTTCTGCCACTCAAGAAGAGTGCAGCGAGCTCGCGCA 1140  
 QY 381 LysLeuValAlaLeuGlyValAsnAlaValAlaTyrTyrArgGlyLeuAspValSerVal 400  
 DB 1141 AMACTGGTCCGCTGGGCGTCAATGCCGTGGCTTACTACCGCGCTTGATGTCCGTC 1200  
 QY 401 IleProThrSerGlyAspValValValAlaThrAlaThrAspAlaLeuMetThrGlyPheThr 420  
 DB 1201 ATCCCGACAGTGGTGACGTGTGCTGGCAACTGACGCCCTCATGACCGCTTTTACC 1260  
 QY 421 GlyAspPheAspSerValIleAspCysAsnThrCysValThrGlnThrValAspPheSer 440  
 DB 1261 GGCGACTTTCGATTCGGTGATAGACTGCAACACGCTGTGTCCACGACAGCTCGACTTCAGC 1320  
 QY 441 LeuAspProThrPheThrIleGluThrIleThrLeuProGlnAspAlaValSerArgThr 460  
 DB 1321 CTTGACCCCTACCTTACCATTGAGACAATCAGCTTCCCGAGATGCTGTCTCCGCTACT 1380  
 QY 461 GlnArgArgGlyArgThrGlyArgGlyLysProGlyIleTyrArgPheValAlaProGly 480  
 DB 1381 CAACGTCGGGGTAGGACTGGCAGAGGAAGCCAGGCATCTACAGATTTGTGGCAGCGGG 1440  
 QY 481 GluArgProSerGlyMetPheAspSerValLeuCysGluCysTyrAspAlaGlyCys 500  
 DB 1441 GAGCGTCTTTCGGCATGTTTGCCTCTGTCTCTGCGAGTGTCTATGACGCGGGTGT 1500  
 QY 501 AlaTyrTyrGluLeuThrProAlaGluThrThrValArgLeuArgAlaTyrMetAsnThr 520  
 DB 1501 GCTTGTATGAGCTTACGCGCGCGGAGACACAGTTAGGCTACGAGCATACATGAACACC 1560  
 QY 521 ProGlyLeuProValCysGlnAspHisLeuGluPheTrpGluGlyValPheThrGlyLeu 540  
 DB 1561 CCGGGACTTCCCGGTGTGCCAAGACCATCTTGAATTTTGGAGGGCGCTCTTTACGGGTCTC 1620  
 QY 541 ThrHisIleAspAlaHisPheLeuSerGlnThrLysGlnSerGlyGluAsnLeuProTyr 560  
 DB 1621 ACCCAGATAGACGCCCACTTCTTATCCAGACAAACAGACAGTGGGGAACCTTCCCTAT 1680  
 QY 561 LeuValAlaTyrGlnAlaThrValCysAlaArgAlaGlnAlaProProProSerTrpAsp 580

Db 1681 CTGTAGCGTACCAAGCCACCGTGTGCGCTAGAGCTCAAGCCCTCCCGCTCGTGGAC 1740  
Qy 581 GlnMetTrpLysCysLeuIleArgLeuLysProThrLeuHisGlyProThrProLeuLeu 600  
Db 1741 CAGATGTGGAAGTCTTGATCGGTCTCAAGCCACCTCCATGGGCCAACCCTCTGCTA 1800  
Qy 601 TyrArgLeuGlyAlaValGlnAsnGluValThrLeuThrHisProValThrLysTyrIle 620  
Db 1801 TATAGACTGGGCGGTGTCCAGAAATGAAGTACCCCTGACGCCACCCAGTCAACAAATATATC 1860  
Qy 621 MetThrCysMetSerAlaAspLeuGluValValThrSerThrTrpValLeuValGlyGly 640  
Db 1861 ATGACATGTATGTGCGCTGACCTGAGGTCTGTCAGAGTACCTGGGTCTGTTGGCGGC 1920  
Qy 641 ValLeuAlaAlaLeuAlaAlaTyrCysLeuSerThrGlyCysValValIleValGlyArg 660  
Db 1921 GTTCTGGCTGTTGGCCGCGTATGCTCTATCCACAGGCTCGGTGCTATAGTAGTAGG 1980  
Qy 661 IleValLeuSerGlyLysProAlaIleIleProAspArgGluValLeuTyrArgGluPhe 680  
Db 1981 ATTGTCTTGTCCGGAAGCCGCAATCATACCCGACAGGAAGTCTCTACCGGGAGTTC 2040  
Qy 681 AspGluMetGluGluCys 686  
Db 2041 GATGAATGGAAGAGTGC 2058

## RESULT 2

AAD31767

ID AAD31767 standard; DNA; 2061 BP.

XX

AC AAD31767;

XX

DT 18-JUN-2002 (first entry)

XX

DE Hepatitis C virus (HCV) NS3/4A DNA coding region.

XX

DE Hepatitis C virus; HCV infection; virucide; fungicide; antibiotic;  
KW cytosstatic; immunostimulant; vaccine; ribavirin; immune response; cancer;  
KW ds.

XX

OS Hepatitis C virus.

XX

FH Key Location/Qualifiers

XX 1. .2061

FT CDS

FT /\*tag= a

FT /product= "HCV NS3/4A protein"

XX

PN WO200213855-A2.

XX

XX 21-FEB-2002.

XX

XX 15-AUG-2001; 2001WO-IB001808.

XX

XX 17-AUG-2000; 2000US-0225767P.

XX

PR 29-AUG-2000; 2000US-0229175P.

XX

PR 03-NOV-2000; 2000US-00705547.

XX

XX (TRIP-) TRIPEP AB.

XX

XX Sallberg M, Hultgren C;

XX

XX WPI; 2002-241837/29.

XX

XX P-PSDB; AAE19900.

XX

XX Vaccine compositions for treating and preventing disease, preferably

XX hepatitis C virus infection, comprises ribavirin and antigen that has

XX epitope present in hepatitis C virus.

XX

XX Claim 1; Page 94-95; 120pp; English.

XX

XX The invention relates to a composition comprising ribavirin and an

XX antigen preferably non structural 3 protein (NS3)/4A fragment of

CC

CC hepatitis C virus (HCV) genome or a peptide or nucleic acid of HCV  
CC sequence. The composition is useful for enhancing an immune response to a  
CC hepatitis C antigen in humans, domestic, sport or pet species and as  
CC vaccines for treating and preventing HCV infections. The composition is  
CC also useful for treating viral, bacterial, fungal diseases and cancer.  
CC The present sequence is HCV NS3/4A DNA coding region

XX  
SQ Sequence 2061 BP; 427 A; 616 C; 571 G; 447 T; 0 U; 0 Other;

## Alignment Scores:

Pred. No.: 3.22e-245 Length: 2061  
Score: 3618.00 Matches: 686  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 6 Gaps: 0

US-09-930-591-2 (1-686) x AAD31767 (1-2061)

Qy 1 MetAlaProIleThrAlaTyrAlaGlnGlnThrArgGlyLeuLeuGlyCysIleIleThr 20  
Db 1 ATGGCGCTATCAGCGCTATGCCAGCAGACAAGGGCCCTTTGGGATGCATAATCACC 60  
Qy 21 SerLeuThrGlyArgAspLysAsnGlnValGlnGluValGlnIleValSerThrAla 40  
Db 61 AGCTTGACCGCGCGGACAAAACCCAGGTGGAGGTGAGGTTCAGATCGTCAACTGCT 120  
Qy 41 AlaGlnThrPheLeuAlaThrCysIleAsnGlyValCysTrpThrValTyrHisGlyAla 60  
Db 121 GCCAGACTTCTTGGCAACTGCTTAACGGGGTGTGGACTGCTTACATGGAGGCC 180  
Qy 61 GlyThrArgThrIleAlaSerProLysGlyProValIleGlnMetTyrThrAsnValAsp 80  
Db 181 GGAACAAGGACCAATTGCTCACCTAAGGTCCTGTTATCCAGATGTACCAATGTGGAC 240  
Qy 81 GlnAspLeuValGlyTrpProAlaProGlnGlyAlaArgSerLeuThrProCysThrCys 100  
Db 241 CAAGACCTCGTAGGCTGGCGCGCTCCCAAGGTGCCCGCTCATTAACACCATGCATTGC 300  
Qy 101 GlySerSerAspLeuTyrLeuValThrArgHisAlaAspValIleProValArgArgArg 120  
Db 301 GGCTCTCGGACCTTTTACCTGGTCACGAGGACCCCGCATTCCTCTGTGCGCGACGG 360  
Qy 121 GlyAspGlyArgGlySerLeuLeuSerProArgProIleSerTyrLeuLysGlySerSer 140  
Db 361 GGTGATGGCAGGGGCGAGCTGCTTTTCGCCCGCGCTATCTCTTACTTGAAGGCTCTCG 420  
Qy 141 GlyGlyProLeuLeuCysProAlaGlyHisAlaValGlyIlePheArgAlaAlaValCys 160  
Db 421 GGAGGCCCTCTGCTGTGCCCGCGAGACATGCCGTAGGCATATTTACAGAGCGCGGTATGC 480  
Qy 161 ThrArgGlyValAlaLysAlaValAspPheIleProValGluSerLeuGluThrThrMet 180  
Db 481 ACCGTGGAGTGGCTAAGGCGGTGGACTTATCCCGGTAGAGAGCTTAGAGACCAACATG 540  
Qy 181 ArgSerProValPheSerAspAsnSerSerProProAlaValProGlnSerTyrGlnVal 200  
Db 541 AGTCCCCGGTGTCTCAGACAACCTCTCCCCACAGCAGAGTCCCGCAGAGTACCAAGTG 600  
Qy 201 AlaHisLeuHisAlaProThrGlySerGlyLysSerThrLysValProAlaAlaTyrAla 220  
Db 601 GCCCACCTGCATGCTCCCGCGGCGGTAAAGACCAAGAGTCCCGCGCGCATACGCA 660  
Qy 221 AlaGlnGlyTyrLysValLeuValLeuAsnProSerValAlaAlaThrMetGlyPheGly 240  
Db 661 GCTCAGGGCTACAGGTGGTGGTCTCAACCCCTCCGTGCTGCAACAATGGGCTTTGGT 720  
Qy 241 AlaTyrMetSerLysAlaHisGlyIleAspProAsnIleArgThrGlyValArgThrIle 260  
Db 721 GCTTACATGTCCAAGGCCCATGGATTGATCTTAACATCAGGACTGGGGTGAGGCAATT 780  
Qy 261 ThrThrGlySerProIleThrTyrSerThrTyrGlyLysPheLeuAlaAspGlyCys 280

Db 781 ACTACTGGCAGCCGATCAGTATTCACCTACCGCAAGTTCCTTGGCGAGCGGGTGT 840  
Qy SerGlyGlyAlaThrAspIleIleCysAspGluCysHisSerThrAspAlaThrSer 300  
Db TCAGGGGGTCTTATGACATATATTTTGGACAGTGCACCTCCACGGATGCACATCC 900  
Qy IleLeuGlyIleGlyThrValLeuAspGlnAlaGluThrAlaGlyAlaArgLeuThrVal 320  
Db ATCTTGGGCATTGGCAGTCTCTTACCAAGCAGAGACCGGGGGGAGACTGACTGTG 960  
Qy LeuAlaThrAlaThrProProGlySerValThrValProHisProAsnIleGluGluVal 340  
Db CTGCCACCGCTACCCCTCGGGCTCGCTCACTGTGCCCTTACATCTACATCGAGGAGTT 1020  
Qy AlaLeuSerThrThrGlyGluIleProPheThrGlyLysAlaIleProLeuGluAlaIle 360  
Db GCTCTGTCCACTACCGAGAGATCCCTTTTATGGCAAGGCTATTTCCCTTGAAGCAAT 1080  
Qy LysGlyGlyArgHisLeuIlePheCysHisSerLysLysCysAspGluLeuAlaAla 380  
Db AAGGGGGGAGACATCTCTCTCTCCACTCAAAGAAAGAGTGGCAGAGCTCGCGCA 1140  
Qy LysLeuValAlaLeuGlyValAsnAlaValAlaThrArgGlyLeuAspValSerVal 400  
Db AAATGTCGTGGTGGCGCTCAATGCGCGTGGCTTACTACCGCGCTTGTATGTGTCGTC 1200  
Qy IleProThrSerGlyAspValValValAlaThrAspAlaLeuMetThrGlyPheThr 420  
Db ATCCGACCAAGTGTGAGTGTGCTGTGGCACTGACCCCTCATGACCGCTTTTACC 1260  
Qy GlyAspPheAspSerValIleAspCysAsnThrCysValThrGlnThrValAspPheSer 440  
Db GCGACCTCGATTCCGTGTAGACTGCAACACGTGTGTACCACGACAGTCCGCTCAGC 1320  
Qy LeuAspProThrPheThrIleGluThrIleThrLeuProGlnAspAlaValSerArgThr 460  
Db CTTGACCTTACCTTACCAITTAGACAAATCAGCTTCCCGAGGATGCTGTCTCCGCTACT 1380  
Qy GlnArgArgGlyArgThrGlyArgGlyLysProGlyIleThrArgPheValAlaProGly 480  
Db CACGTCGGGTAGGACTGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1440  
Qy GluArgProSerGlyMetPheAspSerValLeuCysGluCysThrAspAlaGlyCys 500  
Db GAGCGTCTCTCTGCGATGTTGACTGCTGTCTCTCTGCGAGTGTATGACGCGGTGT 1500  
Qy AlaTrpThrGluLeuThrProAlaGluThrThrValArgLeuArgAlaThrMetAsnThr 520  
Db GCTTGGTATGAGTTAGCGCCCGCGAGACACACAGTTAGGCTACGAGCATACATGAACAC 1560  
Qy ProGlyLeuProValCysGlnAspHisLeuGluPheThrGluGlyValPheThrGlyLeu 540  
Db CCGGACTTCCGTGTGCCAAGACCATCTTGAATTTTGGAGGGCGTCTTTACGGTCTC 1620  
Qy ThrHisIleAspAlaHisPheLeuSerGlnThrLysGlnSerGlyGluAsnLeuProThr 560  
Db ACCACATAGACGCCACTTCTTATCCAGACAAAGCAGAGTGGGAAACCTTCCCTAT 1680  
Qy LeuValAlaThrGlnAlaThrValCysAlaArgAlaGlnAlaProProSerThrAsp 580  
Db CTGGTAGCGTACCAAGCACCGTGTGGCTGTAGAGCTCAAGCCCTCCCGCTCGTGGGAC 1740  
Qy GlnMetTrpLysCysLeuIleArgLeuLysProThrLeuHisGlyProThrProLeuLeu 600  
Db CAGATGGGAAGTGTGTATCTCTTCAAGCCACCTCTCATGGGCCAACACCTCTGCTA 1800  
Qy TyrArgLeuGlyAlaValGlnAsnGluValThrLeuThrHisProValThrLysThrIle 620  
Db TATAGATGGCGGTGTCTCAGATGAATGAGTCACTCCCTGACGACCCAGTACCAAGTATATC 1860  
Qy MetThrCysMetSerAlaAspLeuGluValValThrSerThrTrpValLeuValGly 640  
Db ATGACATGTATGTGGCTGACCTGGAGGTGCTCACGAGTACCTGGGTGCTCGTGGCGC 1920

Qy 641 ValLeuAlaAlaLeuAlaIleValThrCysLeuSerThrGlyCysValValIleValGlyArg 660  
Db 1921 GTTCTGGCTGCTTGGCCGGGTATTGCCCTATCCACAGGCTGGGTGTCATAGTAGGTAGG 1980  
Qy 661 IleValLeuSerGlyLysProAlaIleIleProAspArgGluValLeuThrArgGluPhe 680  
Db 1981 ATTGTCTTGTCCGAAAGCCGCAATCATACCCGACAGGGAAGTCTCTTACCGGAGTTC 2040  
Qy 681 AspGluMetGluGluCys 686  
Db 2041 GATGAATGGAAGAGTGC 2058  
RESULT 3  
AAD60868  
ID AAD60868 standard; DNA; 2061 BP.  
XX  
AC AAD60868;  
DT XX  
XX 15-JAN-2004 (first entry)  
XX Hepatitis C virus NS3/4A DNA.  
XX Ribavirin; vaccine; immune response; infection; therapy; immunostimulant;  
XX virucide; ds.  
XX Hepatitis C virus.  
XX  
FH Key Location/Qualifiers  
FT CDS 1..2061  
FT /\*tag= a  
FT /product= "Hepatitis C virus protein"  
XX  
PN US2002136740-A1.  
XX  
PD 26-SEP-2002.  
XX  
PF 15-AUG-2001; 2001US-00929955.  
XX  
PR 17-AUG-2000; 2000US-0225767P.  
PR 29-AUG-2000; 2000US-0229175P.  
XX  
XX (SALL/) SALLBERG M.  
PA (HULT/) HULTGREN C.  
XX  
PI Sallberg M, Hultgren C;  
XX  
XX WPI; 2003-764978/72.  
DR P-PSDB; ABW00351.  
XX  
PT Vaccine compositions for treating and preventing disease, preferably  
PT hepatitis C virus infection, comprises ribavirin and antigen that has  
PT epitope present in hepatitis C virus.  
XX  
PS Claim 1; Page 60-61; Opp; English.  
XX  
XX The invention relates to a composition comprising ribavirin and an  
CC antigen, where the antigen is derived from a hepatitis virus. The vaccine  
CC is useful in enhancing the immune response to a hepatitis C antigen where  
CC the composition is delivered to an animal identified as requiring an  
CC enhanced immune response. The vaccine is useful in the treatment and  
CC prevention of hepatitis C infection. The present sequence is Hepatitis C  
CC virus NS3/4A DNA  
XX  
SQ Sequence 2061 BP; 427 A; 616 C; 571 G; 447 T; 0 U; 0 Other;

Alignment Scores:  
Pred. No.: 3,22e-245 Length: 2061  
Score: 3618.00 Matches: 686  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 10 Gaps: 0

US-09-930-591-2 (1-686) x AAD60868 (1-2061)

Qy 1 MetAlaProIleThrAlaTyrAlaGlnGlnThrArgGlyLeuLeuGlyCysIleIleThr 20  
Db 1 ATGGCGCTATACGGCTATGCCAGACAGACAGAGGGGCTTTGGGATGCATATCAC 60  
Qy 21 SerLeuThrGlyArgAspLysAsnGlnValGluGlyValGlnIleValSerThrAla 40  
Db 61 AGCTTGACCGCGCGGACAAACACAGGTGGAGGGTGCAGTTCAGTTCGTCACCTGCT 120  
Qy 41 AlaGlnThrPheLeuAlaThrCysIleAsnGlyValCysTrpThrValTyrHisGlyAla 60  
Db 121 GCCCAGACTTTCCTTGGCAACCTGCATTAACGGGGTGTGTGGACTGTCTACATGGAGCC 180  
Qy 61 GlyThrArgThrIleAlaSerProLysGlyProValIleGlnMetTyrThrAsnValAsp 80  
Db 181 GGAACAGACCATTCGCTACCTAAGGGTCTCTGTATCCAGATGTACCAATGTGGAC 240  
Qy 81 GlnAspLeuValGlyTyrProAlaProGlnGlyAlaArgSerLeuThrProCysThrCys 100  
Db 241 CAAGACCTCGTAGCTGGCCCGCTCCCAAGGTGCGCGCTCATTAACACCATGCACTGC 300  
Qy 101 GlySerSerAspLeuTyrLeuValThrArgHisAlaAspValIleProValArgArg 120  
Db 301 GGCTCCTCGGACCTTACCTGGTACGAGGCACCGCGATGTCATTCCTGTGCGCGGACGG 360  
Qy 121 GlyAspGlyArgGlySerLeuLeuSerProArgProIleSerTyrLeuLysGlySer 140  
Db 361 GGTGATGGCAGGGGCGCTGCTTCGCCCGGCTATCTCTTACTTGAAGGCTCTCTCG 420  
Qy 141 GlyGlyProLeuLeuCysProAlaGlyHisAlaValGlyIlePheArgAlaAlaValCys 160  
Db 421 GGAGGCTCTGCTGTGCCCCCGCAGGACATGCCGTAGGCATATTACAGAGCGCGGTATGC 480  
Qy 161 ThrArgGlyValAlaLysAlaValAspPheIleProValGluSerLeuGluThrThrMet 180  
Db 481 ACCGTGGAGTGGGTGAAGCGGTGGACTTCATCCCGGTAGAGCTTAGACACAACCATG 540  
Qy 181 ArgSerProValPheSerAspAsnSerSerProProAlaValProGlnSerTyrGlnVal 200  
Db 541 AGTCCCCGGGTGCTTCAGACACTCTCTCCACACAGCAGTGCCTCCAGAGTCAACAGTG 600  
Qy 201 AlaHisLeuHisAlaProThrGlySerGlyLysSerThrLysValProAlaAlaTyrAla 220  
Db 601 GCCCACCTGCATGCTCCACCGCAGCGGTAAAGACCAAGGTCCCGCGCGCATACGCA 660  
Qy 221 AlaGlnGlyTyrLysValLeuValLeuAsnProSerValAlaAlaThrMetGlyPheGly 240  
Db 661 GCTCAGGGCTACAAGGTGCTGCTCAACCCCTCCGTTGCTGCAACAATGGGCTTTGGT 720  
Qy 241 AlaTyrMetSerLysAlaHisGlyIleAspProAsnIleArgThrGlyValArgThrIle 260  
Db 721 GCTTACATGTCACAGGCCCATGGGATGATCCTAACATCAGAGCTGGGGTGAGACAAAT 780  
Qy 261 ThrThrGlySerProIleThrTyrSerThrTyrGlyLysPheLeuAlaAspGlyGlyCys 280  
Db 781 ACTACTGCACCCCGATACGTATTCCACCTACCGCAAGTTCCTTGGCGCGGGGTGT 840  
Qy 281 SerGlyGlyAlaTyrAspIleIleCysAspGluCysHisSerThrAspAlaThrSer 300  
Db 841 TCAGGGGTGCTTATGACATAATAATTGTGACGAGTGCCTCCACCGATGCAACATCC 900  
Qy 301 IleLeuGlyIleGlyThrValLeuAspGlnAlaGluThrAlaGlyAlaArgLeuThrVal 320  
Db 901 ATCTTGGGCATGGCACTGCTCTTGACCAAGCAGACCGCGGGGGCGAGACTGACTGTG 960  
Qy 321 LeuAlaThrAlaThrProGlySerValThrValProHisProAsnIleGluGluVal 340  
Db 961 CTGCGCACCGTACCCCTCCGGGCTCCGTGCTGCTGCCCCCATCTTAACATCGAGAGGTT 1020  
Qy 341 AlaLeuSerThrThrGlyGluIleProPheTyrGlyLysAlaIleProLeuGluAlaIle 360  
|||||

RESULT 4  
ADG47658  
ID ADG47658 standard; DNA; 2061 BP.

Db 1021 GCTCTGTCCACTACCGGAGAGATCCCTTTTATGGCAAGGCTATTCCCTTGAAGCAATT 1080  
Qy 361 LysGlyGlyArgHisLeuIlePheCysHisSerLysLysCysAspGluLeuAlaAla 380  
Db 1081 AAGGGGGGAGACATCTCACTCTCTGCACTCAAAAGAAAGTGCAGACGAGCTCGCGCA 1140  
Qy 381 LysLeuValAlaLeuGlyValAsnAlaValAlaTyrTyrArgGlyLeuAspValSerVal 400  
Db 1141 AAATGTGTCGCTTGGCGCTCAATGCGGTGCTTACTACCGCGCTTGTATGTCTCGTCTC 1200  
Qy 401 IleProThrSerGlyAspValValValValAlaThrAspAlaLeuMetThrGlyPheThr 420  
Db 1201 ATCCCGACAGTGTGTGAGCTTGTCTGTCGCAACTACGCCCTCATGACCGGCTTTACC 1260  
Qy 421 GlyAspPheAspSerValIleAspCysAsnThrCysValThrGlnThrValAspPheSer 440  
Db 1261 GCGGACTTCGATTCCGTTGATAGACTGCAACACAGTGTGTCAACCCAGACAGTCTCAGC 1320  
Qy 441 LeuAspProThrPheThrIleGluThrIleThrLeuProGlnAspAlaValSerArgThr 460  
Db 1321 CTTGACCTTACCTTACCATTTGAGACAATCAGCTTCCCGAGGATGCTCTCCCGTACT 1380  
Qy 461 GlnArgArgGlyArgThrGlyArgGlyLysProGlyIleTyrArgPheValAlaProGly 480  
Db 1381 CAACGTCCGGGTAGGACTGGCAGAGGAAGCCAGGATCTACAGATTTGTGGCACCGGG 1440  
Qy 481 GluArgProSerGlyMetPheAspSerSerValLeuCysGluCysTyrAspAlaGlyCys 500  
Db 1441 GAGGCTCTTCTGCGATGTTGACTCGTCTCTCTCTGCGAGTGTATGACGCGGGTGT 1500  
Qy 501 AlaTyrTyrGluLeuThrProAlaGluThrThrValArgLeuArgAlaTyrMetAsnThr 520  
Db 1501 GCTTGGTATGAGCTTACGCCCGCGCAGACCACTAGGTACGAGCATACATGAACACC 1560  
Qy 521 ProGlyLeuProValCysGlnAspHisLeuGluPheTrpGluGlyValPheThrGlyLeu 540  
Db 1561 CCGGGACTTCCCGTGTGCCAAGACCCTTGAATTTTGGGAGGGCGTCTTTACGGGCTC 1620  
Qy 541 ThrHisIleAspAlaHisPheLeuSerGlnThrLysGlnSerGlyGluAsnLeuProTyr 560  
Db 1621 ACCACATAGACGCCCATCTTCTATCCAGACAAGCAGAGTGGGAAACCTTCCCTAT 1680  
Qy 561 LeuValAlaTyrGlnAlaThrValCysAlaArgAlaGlnAlaProProProSerTrpAsp 580  
Db 1681 CTGCTAGCGTACCAAGCCACCGTGTGCGGTAGAGCTCAAGCCCTCCCGCTCGTGGGAC 1740  
Qy 581 GlnMetTrpLysCysLeuIleArgLeuLysProThrLeuHisGlyProThrProLeuLeu 600  
Db 1741 CAGATGGGAAGTCTTGTATCCCGTCTCAAGCCCACTCCATGGGCCCAACCTCTGCTA 1800  
Qy 601 TyrArgLeuGlyAlaValGlnAsnGluValThrLeuThrHisProValThrLysTyrIle 620  
Db 1801 TATAGCTGGGCGCTGTCCAGATGAAGTCACTCCCTGACGACCCAGTCAACATATATC 1860  
Qy 621 MetThrCysMetSerAlaAspLeuGluValValThrSerThrTrpValLeuValGlyGly 640  
Db 1861 ATGACATGTATGTGCGGTGACCTGGAGTGTGTACAGAGTACCTGGGTGCTCGTGGCGGC 1920  
Qy 641 ValLeuAlaAlaLeuAlaAlaTyrCysLeuSerThrGlyCysValValIleValGlyArg 660  
Db 1921 GTTCTGCTCTTGGCGCGGTATGTGCTATCCACAGGCTGCGGTGCTATAGTAGGAGG 1980  
Qy 661 IleValLeuSerGlyLysProAlaIleIleProAspArgGluValLeuTyrArgGluPhe 680  
Db 1981 ATTGTCTTGGCGAAAGCCGGCAATCATCCCGACAGGAGAGTCTCTACCGGGAGTTC 2040  
Qy 681 AspGluMetGluGluCys 686  
|||||  
Db 2041 GATGAATGGAAGAGTGC 2058  
|||||

XX AC ADG47658;  
 XX DT 11-MAR-2004 (first entry)  
 XX DE HCV NS3/4A domain DNA.  
 XX KW ds; gene; immunogen; hepatitis C virus; HCV infection; vaccine.  
 XX OS Hepatitis C virus.  
 XX FH Key Location/Qualifiers  
 XX FT CDS 1..2061  
 XX FT /\*tag= a  
 XX FT /product= "NS3/4A domain"  
 XX US2003206919-A1.  
 XX PN 06-NOV-2003.  
 XX PD 26-NOV-2002; 2002US-00307047.  
 XX PF 17-AUG-2000; 2000US-0225767P.  
 XX PR 29-AUG-2000; 2000US-0229175P.  
 XX PR 15-AUG-2001; 2001US-00929955.  
 XX PR 15-AUG-2001; 2001US-00930591.  
 XX PA (SALL/) SALLBERG M.  
 XX PI Sallberg M;  
 XX PI WPI; 2004-051480/05.  
 XX DR P-PSDB; ADG47659.  
 XX DR  
 XX PT New purified or isolated nucleic acid useful for enhancing an immune  
 XX PT response to a hepatitis C antigen comprises specific nucleotide sequences  
 XX PT and the amino acid sequences.  
 XX PS Example 1; SEQ ID NO 1; 83pp; English.  
 XX CC The invention relates to a purified or isolated nucleic acid. The  
 XX CC peptides are useful as immunogens for the treatment and prevention of  
 XX CC hepatitis C virus (HCV) infection, in vaccine and immunogen compositions.  
 XX CC The nucleic acid and the peptide enhance an immune response to a  
 XX CC hepatitis C antigen and are potent immunogens. The present sequence is  
 XX CC used in the exemplification of the invention.  
 XX SQ Sequence 2061 BP; 427 A; 616 C; 571 G; 447 T; 0 U; 0 Other;  
 Alignment Scores:  
 Pred. No.: 3 22e-245 Length: 2061  
 Score: 3618.00 Matches: 686  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 12 Gaps: 0  
 US-09-930-591-2 (1-686) x ADG47658 (1-2061)  
 QY 1 MetAlaProIleThrAlaTyrAlaGlnThrArgGlyLeuLeuGlyCysIleIleThr 20  
 DB 1 ATGGCGCTATACGGCTATGCCAGCAGACAAAGGGCCCTTTGGATGCATAATCACC 60  
 QY 21 SerIleuThrGlyArgAspLysAenGlnValGluGlyValGlnIleValSerThrAla 40  
 DB 61 AGCTTGACCGCGCGGACAAAACCAAGGTGGAGGTGAGGTTCAGATCGGTCAACTGCT 120  
 QY 41 AlaGlnThrPheLeuAlaThrCysIleAenGlyValCysTrpThrValTyrHisGlyAla 60  
 DB 121 GCCCAGACTTCTTGGCAACTGCATTAACGGGTGTGTGGACTGTCTACCATGGAGCC 180  
 QY 61 GlyThrArgThrIleAlaSerProLysGlyProValIleGlnMetTyrThrAsnValAsp 80

DB 181 GGAACAAGACCATTGCGTCACTAAGGGTCTCTTATCCAGATGTACACCAATGTGGAC 240  
 QY 81 GlnAspLeuValGlyTyrProAlaProGlnGlyAlaArgSerLeuThrProCysThrCys 100  
 DB 241 CAAGACCTCGTAGGCTGGCCCGCTCCCAAGGTGCCCGCTCATTAACACATGCACTTGC 300  
 QY 101 GlySerSerAspLeuTyrLeuValThrArgHisAlaAspValIleProValArgArgArg 120  
 DB 301 GGCTCTCTCGACCTTTACCTGTGTACAGAGCAGCGCGATGTATCTCTCTGTGGCCGACGG 360  
 QY 121 GlyAspGlyArgGlySerLeuLeuSerProArgProIleSerTyrLeuLysGlySerSer 140  
 DB 361 GGTGATGGCAGGGCAGCCTGCTTTCGCCCGCGCTATCTCTTACTTGAAGGCTCCTCG 420  
 QY 141 GlyGlyProLeuLeuCysProAlaGlyHisAlaValGlyIlePheArgAlaAlaValCys 160  
 DB 421 GGAGGCCCTCTGCTGTGCCCGCAGACATGCCGTAGGCATATTTCAGACCGCGGTATGC 480  
 QY 161 ThrArgGlyValAlaLysAlaValAspPheIleProValGluSerLeuGluThrThrMet 180  
 DB 481 ACCGTGGAGTGGCTAAGCGGTGACTTCATCCCGTAGAGAGCTTAGACACACCATG 540  
 QY 181 ArgSerProValPheSerAspAsnSerSerProAlaValProGlnSerTyrGlnVal 200  
 DB 541 AGGTCCCGGTGTCTCTCAGACAACTCTCCACCAGCAGTGCCTCCAGAGCTACCAAGTG 600  
 QY 201 AlaHisLeuHisAlaProThrGlySerGlyLysSerThrLysValProAlaAlaTyrAla 220  
 DB 601 GCCCAGCTGCATGCTCCACCGCAGCGGTAGAGCACCAGGTCCCGCCGCATACGCA 660  
 QY 221 AlaGlnGlyTyrLysValLeuValLeuAnProSerValAlaAlaThrMetGlyPheGly 240  
 DB 661 GCTCAGGGCTACAGGTGTGTGTCTCAACCTCCGTGCTGCAACATGGCTTGTGT 720  
 QY 241 AlaTyrMetSerLysAlaHisGlyIleAspProAenIleArgThrGlyValArgThrIle 260  
 DB 721 GCTTACATGTCGAAGGCCCATGGCATGTATCTTAACATCAGGACTGGGTGAGGACAT 780  
 QY 261 ThrThrGlySerProIleThrTyrSerThrTyrGlyLysPheLeuAlaAspGlyGlyCys 280  
 DB 781 ACTACTGGCAGCGCCGATACGATTCACCTACGGCAAGTTCCTTGCCACGCGGGGTGT 840  
 QY 281 SerGlyGlyValTyrAspIleIleCysAspGluCysHisSerThrAspAlaThrSer 300  
 DB 841 TCAGGGGTGTCTATGACATATATTTGTGACGAGTGCCACTCCACGATGCAACATCC 900  
 QY 301 IleLeuGlyIleGlyThrValLeuAspGlnAlaGluThrAlaGlyAlaArgIleThrVal 320  
 DB 901 ATCTTGGCATTTGGCACTGCTCTTGACCAAGCAGAGACCGCGGGGCGAGACTGACTGTG 960  
 QY 321 LeuAlaThrAlaThrProGlySerValThrValProHisProAsnIleGluGluVal 340  
 DB 961 CTCGCCACCGCTACCCCTCCGGGCTCCGTCATCTGTGCCCATCTTAACATCAGGAGGT 1020  
 QY 341 AlaLeuSerThrThrGlyGluIleProPheTyrGlyLysAlaIleProLeuGluAlaIle 360  
 DB 1021 GCTCTGTCCACTACCGGAGAGATCCCTTTTATGCCAAGGCTATTCCTCTTGAAGCAAT 1080  
 QY 361 LysGlyGlyArgHisLeuIlePheCysHisSerLysLysCysAspGluIleAlaAla 380  
 DB 1081 AAGGGGGGAGACATCTCATCTTCTGCCACTCAAAAGAGAGTGGCAGAGCTCGCGCA 1140  
 QY 381 LysLeuValAlaLeuGlyValAsnAlaValAlaTyrTyrArgGlyLeuAspValSerVal 400  
 DB 1141 AAACGTGTGCGGTGGCGCTCAATGCGGTGCTTACTACCGGCGCTTGTATGTCCGTC 1200  
 QY 401 IleProThrSerGlyAspValValValAlaThrAspAlaLeuMetThrGlyPheThr 420  
 DB 1201 ATCCCGACAGTGTGTGCTGTGTGGCACTGACGCCCTCATGCCGCTTTACC 1260  
 QY 421 GlyAspPheAspSerValIleAspCysAsnThrCysValThrGlnThrValAspPheSer 440  
 DB 1261 GCGGACTTCGATTGCGTGTAGACTGCAACACGTGTGTCCACGACAGTCAGCTTCAGC 1320

Qy 441 LeuAspProThrPheThrIleGluThrIleThrLeuProGlnAspAlaValSerArgThr 460  
Db 1321 CTTGACCCCTACCTTACCATTGAGACAATCAGCTTCCAGGATGCTGTCTCCGGTACT 1380  
Qy 461 GlnArgArgGlyArgThrGlyArgGlyPProGlyIleThrArgPheValAlaProGly 480  
Db 1381 CAACGTCGGGGTAGGACTGGCAGAGGGAAGCAGGCATCTACAGATTTGTGGCACCGGGG 1440  
Qy 481 GluArgProSerGlyMetPheAspSerValLeuLeuGluCysGluCysThrAspAlaGlyCys 500  
Db 1441 GAGCGTCTCTGGCATGTTGACTCGTCTGCTCTCGGAGTGTATGACGCGGGTGT 1500  
Qy 501 AlaTrpThrGluLeuThrProAlaGluThrThrValArgLeuArgAlaTyrMetAsnThr 520  
Db 1501 GCTTGGTATGACTTACGCCCGCCGAGACCAAGTACGAGTACGAGCATACTGAACACC 1560  
Qy 521 ProGlyLeuProValCysGlnAspHisLeuGluPheTrpGluGlyValPheThrGlyLeu 540  
Db 1561 CCGGACTTCCGGTGTGCCAAGACCATCTTGAATTTTGGGAGGGCGTCTTTACGGGTCTC 1620  
Qy 541 ThrHisIleAspAlaHisPheLeuSerGlnThrLysGlnSerGlyGluAsnLeuProTyr 560  
Db 1621 ACCACATAGAGCCCATCTCTATCCAGACAAGACAGTGGGAAAACCTTCCCTAT 1680  
Qy 561 LeuValAlaTyrGlnAlaThrValCysAlaArgAlaGlnAlaProProSerTrpAsp 580  
Db 1681 CTGGTAGCTACCAAGCCACCGTGTGCGCTAGAGCTCAAGCCCTCCCTCCCGTGTGGGAC 1740  
Qy 581 GlnMetTrpLysCysLeuIleArgLeuLysProThrLeuHisGlyProThrProLeuLeu 600  
Db 1741 CAGATGTGGAAGTGTGTCGCTCAAGGCCACCTCCATGGGCCCAACACCTCTGCTA 1800  
Qy 601 TyrArgLeuGlyAlaValGlnAsnGluValThrLeuThrHisProValThrLysTyrIle 620  
Db 1801 TATGACTGGCGCTGTCCAGATGAGTACCTGAGCGACCCAGTCCACAGTATATC 1860  
Qy 621 MetThrCysMetSerAlaAspLeuGluValValThrSerThrTrpValLeuValGlyGly 640  
Db 1861 ATGACATGATGTGCGCTGACCTGGAGGTGTGTCAGAGTACCTGGGTGCTCTGTGGCGGC 1920  
Qy 641 ValLeuAlaAlaLeuAlaAlaTyrCysLeuSerThrGlyCysValValIleValGlyArg 660  
Db 1921 GTTCTGGCTGTCTGGCGCGTATTGCTATCCACAGGCTGCGTGTGTCATAGTAGTAGG 1980  
Qy 661 IleValLeuSerGlyLysProAlaTyrIleProAspArgGluValLeuTyrArgGluPhe 680  
Db 1981 ATTGTCTTGTCCGGAAGCCGGCAATCATACCCGACAGGGAAGTCTCTACCGGGAGTTC 2040  
Qy 681 AspGluMetGluGluCys 686  
Db 2041 GATGAATGGAGAGTGC 2058

## RESULT 5

ADG47692

ID ADG47692 standard; DNA; 2078 BP.

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

PD 06-NOV-2003.  
XX 26-NOV-2002; 2002US-00307047.  
XX 17-AUG-2000; 2000US-0225767P.  
XX 29-AUG-2000; 2000US-0229175P.  
XX 15-AUG-2001; 2001US-00929955.  
XX 15-AUG-2001; 2001US-00930591.  
XX (SALL/) SALLBERG M.  
XX Sallberg M;  
XX WPI; 2004-051480/05.  
XX P-PSDB; ADG47693.  
XX New purified or isolated nucleic acid useful for enhancing an immune response to a hepatitis C antigen comprises specific nucleotide sequences and the amino acid sequences.  
XX Claim 2; SEQ ID NO 35; 83pp; English.  
XX The invention relates to a purified or isolated nucleic acid. The peptides are useful as immunogens for the treatment and prevention of hepatitis C virus (HCV) infection, in vaccine and immunogen compositions. The nucleic acid and the peptide enhance an immune response to a hepatitis C antigen and are potent immunogens. The present sequence is used in the exemplification of the invention.  
XX SQ Sequence 2078 BP; 351 A; 834 C; 630 G; 263 T; 0 U; 0 Other;

## Alignment Scores:

Pred. No.: 3,26e-245 Length: 2078  
Score: 3618.00 Matches: 686  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 12 Gaps: 0

US-09-930-591-2 (1-686) x ADG47692 (1-2078)

Qy 1 MetAlaProIleThrAlaTyrAlaGlnGlnThrArgGlyLeuLeuGlyCysIleThr 20  
Db 12 ATGGCCCATCATCGGCTACGCCAGACAGCCGCGCTGTGGGTGTCATCATCACC 71  
Qy 21 SerLeuThrGlyArgAspLysAsnGlnValGluGlyGluValGlnIleValSerThAla 40  
Db 72 AGCTGTACCGCGCGCCGCAAGAACCCAGGTGGAGGGCGAGGTGCAGATCGTGAGCACCGCC 131  
Qy 41 AlaGlnThrPheLeuAlaThrCysIleAsnGlyValCysTrpThrValTyrHisGlyAla 60  
Db 132 GCCAGACCTTCTTGGCCACCTGCATCAACGGCGTGTGTGGACCGGTATACACGGCGCC 191  
Qy 61 GlyThrArgThrIleAlaSerProLysGlyProValIleGlnMetTyrThrAsnValAsp 80  
Db 192 GGCACCGCACCATCGCCAGCCGCCAAGGGCCCGCTGATCCAGATGTACCAACAGTGGAC 251  
Qy 81 GlnAspLeuValGlyTrpProAlaProGlnGlnAlaArgSerLeuThrProCysThrCys 100  
Db 252 CAGGACCTGTGGTGGCTGGCCCGCCCGCCAGGGCGCGCCAGCTGACCCCTGTACACCTGC 311  
Qy 101 GlySerSerAspLeuTyrLeuValThrArgHisAlaAspValIleProValArgArg 120  
Db 312 GGCAGACGACCTGTACCTGTGTGACCCGCGCAGCGGCGAGTGTATCCCGTGGCGCGCGC 371  
Qy 121 GlyAspGlyArgGlySerLeuLeuSerProArgProIleSerTyrLeuLysGlySerSer 140  
Db 372 GGCAGCGCGCGCGAGCTGTGTGAGCCCGCCCGCCCATCAGTACCTGAGAGGGCAGCAGC 431  
Qy 141 GlyGlyProLeuLeuCysProAlaGlyHisAlaValGlyIlePheArgAlaAlaValCys 160  
Db 432 GGCAGCGCGCGCGTGTGTGCGCCCGCGCGCACCGGTGGGCATCTTCCGCGCGCGCGTGTGC 491

```
QY 161 ThrArgGlyValAlaLysAlaValAspPheIleProValGluSerLeuGluThrThrMet 180
DB 492 ACCGCGCGGTGGCCAGGCGGTGACTTATCCCTGGAGAGCTGGAGACCACTATG 551
QY 181 ArgSerProValPheSerAspAsnSerSerProProAlaValProGlnSerTyrGlnVal 200
DB 552 CGCAGCCCGGTGTTACGCGCAACAGCAGCGCCCGCCCGCGCGTGCACAGACTACCAAGTG 611
QY 201 AlaHisLeuHisAlaProThrGlySerGlyLysSerThrLysValProAlaAlaTyrAla 220
DB 612 GCCCACTGACGCGCCCGCCAGCGGCAAGAGCACCAGGTGCGCCGCGCTACGCC 671
QY 221 AlaGlnGlyTyrLysValLeuValLeuAsnProSerValAlaAlaThrMetGlyPheGly 240
DB 672 GCCCAGGGCTACAAAGTGTGTGCTGAACCCCGAGCGTGGCCGCCACCATGGCTTCGGC 731
QY 241 AlaTyrMetSerLysAlaHisGlyIleAspProAsnIleArgThrGlyValArgThrIle 260
DB 732 GCCTACATGACCAAGGCCCGCCAGCGCATCGACCCCAACATCCGCGCGGTGCGCACCATC 791
QY 261 ThrThrGlySerProIleThrTyrSerThrTyrGlyLysPheLeuAlaAspGlyGlyCys 280
DB 792 ACACCGCGAGCCCATCACTACAGCACTACGGCAAGTTCTTGGCCGACGGCGGTGC 851
QY 281 SerGlyGlyAlaTyrAspIleIleIleCysAspGluCysHisSerThrAspAlaThrSer 300
DB 852 AGCGCGCGCGCTACGACATCATCTGCGAGGTGCCACAGCAGCCGCGCACCGAC 911
QY 301 IleLeuGlyIleGlyThrValLeuAspGlnAlaGluThrAlaGlyAlaArgLeuThrVal 320
DB 912 ATCCTGGGCATCGGCACCGTGTGGACAGCGCGAGACCGCGCGCGCTGACCGGTG 971
QY 321 LeuAlaThrAlaThrProProGlySerValThrValProHisProAsnIleGluGluVal 340
DB 972 CTGGCCACCGCCACCCCGCGGAGGTGACCGTGCCTGCCCGCCACCCCAACATCGAGGGGTG 1031
QY 341 AlaLeuSerThrThrGlyGluIleProPheTyrGlyLysAlaIleProLeuGluAlaIle 360
DB 1032 GCCTGAGCACCACCGCGGAGATCCCTTCTACGGCAAGGCCATCCCTCTGGAGGCCATC 1091
QY 361 LysGlyGlyArgHisLeuIlePheCysHisSerLysLysCysAspGluLeuAlaAla 380
DB 1092 AAGGGCGCGCGCCACCTGATCTTCTGCCACAGCAAGAAGTGCAGCAGCTGGCGCGC 1151
QY 381 LysLeuValAlaLeuGlyValAsnAlaValAlaTyrTyrArgGlyLeuAspValSerVal 400
DB 1152 AAGCTGTGTCCCTTGGCGGTGAACCGCGTGCCTACTACCGCGCGCTGGACGTGACGGTG 1211
QY 401 IleProThrSerGlyAspValValValAlaThrAspAlaLeuMetThrGlyPheThr 420
DB 1212 ATCCCGCAGCGCGGACGTGTGTGTGTGGCCACCGAGCGCTGATGACCGGCTTCACC 1271
QY 421 GlyAspPheAspSerValIleAspCysAsnThrCysValThrGlnThrValAspPheSer 440
DB 1272 GCGGACTTCGACAGCGGTGATCGATCGCAACACCTGCGTGACCCAGACCGCTGGACTTCAGC 1331
QY 441 LeuAspProThrPheThrIleGluThrIleThrLeuProGlnAspAlaValSerArgThr 460
DB 1332 CTGGACCCCACTTACCATCGAGACCATCACTCCCTGCCCGAGACCGCGTGGCCGACCC 1391
QY 461 GlnArgGlyArgThrGlyArgGlyLysProGlyIleTyrArgPheValAlaProGly 480
DB 1392 CAGCGCGCGCGCGCAGCGCGCGGCAAGCCCGGCATCTACCGCTTCGTGGCGCCCGCGC 1451
QY 481 GluArgProSerGlyMetPheAspSerSerValLeuCysGluCysTyrAspAlaGlyCys 500
DB 1452 GAGCGCGCCCGCGCGCATGTTTCGACAGCAGCGGTGCTGTGGAGTGTACGACCGCGCGTGC 1511
QY 501 AlaTyrPyrGluLeuThrProAlaGluThrThrValArgLeuArgAlaTyrMetAsnThr 520
DB 1512 GCCTGGTACGAGTGCACCCCGCGGAGACCACTCCGCGCGCTGCGCGCTACATGAACACC 1571
QY 521 ProGlyLeuProValCysGlnAspHisLeuGluPheTrpGluGlyValPheThrGlyLeu 540
```

```
DB 1572 CCGGCGCTGCGCGGTGGCCAGGACCACTGGAGTCTGGAGGGCGGTTCACCGCGCTG 1631
QY 541 ThrHisIleAspAlaHisPheLeuSerGlnThrLysGlnSerGlyGluAsnLeuProTyr 560
DB 1632 ACCCATCGAGCGCCCACTTCTCGAGCCAGACCAAGCAGAGCGCGGAGAACCTGCGCTAC 1691
QY 561 LeuValAlaTyrGlnAlaThrValCysAlaArgAlaGlnAlaProProProSerTrpAsp 580
DB 1692 CTGGTGGCTTACAGGCGCACCGTGTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGAC 1751
QY 581 GlnMetTrpLysCysLeuIleArgLeuLysProThrLeuHisGlyProThrProLeuLeu 600
DB 1752 CAGATGTGGAAGTGCCTGATCCGCTGGAAGCCACCGCTGCACGGCGCCACCGCGCTGCTG 1811
QY 601 TyrArgLeuGlyAlaValGlnAsnGluValThrLeuThrHisProValThrLysTyrIle 620
DB 1812 TACCGCGCTGGCGCGCGTGCAGAACAGAGGTGACCTTACCCCGCGCGCGCGCGCGCATC 1871
QY 621 MetThrCysMetSerAlaAspLeuGluValValThrSerThrTrpValLeuValGlyGly 640
DB 1872 ATGACCTGATGAGCGCGCGCTGAGGTGGTGCACAGCAGCTGGGTGCTGGTGGCGCGC 1931
QY 641 ValLeuAlaAlaLeuAlaAlaTyrCysLeuSerThrThrGlyCysValIleValIleValArg 660
DB 1932 GTGCTGGCGCGCTGGCGCGCTACTGCTGAGCAGCGCGCTGCTGCTGATCGTGGCGCGC 1991
QY 661 IleValLeuSerGlyLysProAlaIleIleProAspArgGluValLeuTyrArgGluPhe 680
DB 1992 ATCGTGTCTGAGCGGCAAGCGCGCATCATCCCGCGAGGTGCTGTACCGCGAGTTTC 2051
QY 681 AspGluMetGluGluCys 686
DB 2052 GACGAGATCGAGAGTGC 2069
RESULT 6
AAN92106
ID AAN92106 standard; DNA; 7310 BP.
XX AC AAN92106;
XX DT 09-SEP-2004 (revised)
XX DT 25-MAR-2003 (revised)
XX DT 02-MAR-1990 (first entry)
XX DE Combined ORFs of the HCV cDNAs from clones K9-1 through 15e.
XX KW Hepatitis C virus; HCV; non-A, non-B hepatitis; NANBH; ss.
XX OS Hepatitis C virus.
XX OS Unidentified.
XX FH Key
XX FT CDS Location/Qualifiers
XX FT 3..7310
XX FT /*tag= a
XX EP18216-A.
XX PD 31-MAY-1989.
XX PF 18-NOV-1988; 88EP-00310922.
XX PR 18-NOV-1987; 87US-00122714.
XX PR 30-DEC-1987; 87US-00139886.
XX PR 26-FEB-1988; 88US-00161072.
XX PR 06-MAY-1988; 88US-00191263.
XX PR 26-OCT-1988; 88US-00263584.
XX PR 14-NOV-1988; 88US-00271450.
XX PA (CHIR ) CHIRON CORP.
XX PA (CHIR ) CHIRON CORP.
XX PI Houghton M, Choo QL, Kuo G;
```



XX WPI; 1989-159274/22.  
DR P-PSDB; AAP92050.  
XX Purified hepatitis C virus - and associated nucleic acids and  
PT polypeptide(s).  
XX Claim 3; Fig 47-1 - 47-8; 139pp; English.  
XX It is a double-stranded nucleotide sequence of the open reading frame  
CC (ORF) (tag a) extending through clones K9-1 to 15e of hepatitis C virus  
CC (HCV) cDNA. It can be used to make oligomeric DNA hybridisation probes to  
CC detect the presence of HCV nucleic acids in samples. The polypeptide(s)  
CC it encodes could be used as immunoassay reagents and vaccines and to  
CC generate antibodies useful in diagnosis and passive immunotherapy for HCV  
CC infection/non-A, non-B hepatitis. (Updated on 25-MAR-2003 to correct PR  
CC field.) (Updated on 25-MAR-2003 to correct PI field.)  
CC Revised record issued on 09-SEP-2004 : Correction to keywords  
XX  
SQ Sequence 7310 BP; 1491 A; 2217 C; 2058 G; 1540 T; 0 U; 4 Other;  
Alignment Scores:  
Pred. No.: 1,88e-241 Length: 7310  
Score: 3574.00 Matches: 672  
Percent Similarity: 99.42% Conservative: 10  
Best Local Similarity: 97.96% Mismatches: 4  
Query Match: 98.78% Indels: 0  
DB: 1 Gaps: 0  
US-09-930-591-2 (1-686) x AAN92106 (1-7310)  
QY 1 MetAlaProIleThrAlaTyrAlaGlnGlnThrArgGlyLeuLeuGlyCysIleIleThr 20  
DB 1728 CTGGCGCCCATCAGCGCGTAGCCCGACGACAGACAGAGGGCCCTCTAGGTGTCATATACCC 1787  
QY 21 SerLeuThrGlyArgAspLysAsnGlnValGluGlyGluValGlnIleValSerThrAla 40  
DB 1788 AGCCTAACTGGCCGGGACAAAACCAAGTGAGGGTGAGGTCCAGATTGTCTCAACTGCT 1847  
QY 41 AlaGlnThrPheLeuAlaThrCysIleAsnGlyValCysTyrThrValTyrHisGlyAla 60  
DB 1848 GCCCAAACTTCTTGGCAACGTGCATCAATGGGGTGTCTGGACTGTCTACACGGGGCC 1907  
QY 61 GlyThrArgThrIleAlaSerProLysGlyProValIleGlnMetTyrThrAsnValAsp 80  
DB 1908 GGAACGAGGACATCGCTGACCCAGAGGTCTGTCTATCCAGATGTATACCAATGTAGAC 1967  
QY 81 GlnAspLeuValGlyTyrProAlaProGlnGlyAlaArgSerLeuThrProCysThrCys 100  
DB 1968 CAAGACCTTGTGGGTGGCCCGCTCCGCAAGGTAGCCGCTCATTGACACCCCTGCACCTTGC 2027  
QY 101 GlySerSerAspLeuTyrLeuValThrArgHisAlaAspValIleProValArgArg 120  
DB 2028 GGCTCTCGGACCTTATCTGGTCACGAGGACGCGGATGTCTTCCGTGCGCGCGG 2087  
QY 121 GlyAspGlyArgGlySerLeuLeuSerProArgProIleSerTyrLeuLysGlySerSer 140  
DB 2088 GGTGATACGAGGCGGAGCTGCTGTCCCGGCCCATTTCTACTTGAAGGCTCTCTCG 2147  
QY 141 GlyGlyProLeuLeuCysProAlaGlyHisAlaValGlyIlePheArgAlaAlaValCys 160  
DB 2148 GGGGTCTCGCTGTGTGCCCCCGGGGCACGCCGTGGGCATATTAGGGCCGCGGTGTC 2207  
QY 161 ThrArgGlyValAlaLysAlaValAspPheIleProValGluSerLeuGluThrThrMet 180  
DB 2208 ACCGTGGAGTGGCTAAGCGGTGACTTTATCCCTGTGGGAACCTTAGAGACAACCATG 2267  
QY 181 ArgSerProValPheSerAspAsnSerSerProAlaValProGlnSerTyrGlnVal 200  
DB 2268 AGGTCCCGGTGTTCCAGCGTAACCTCTCTCCACGATGAGTGGCCCGAGGTCTCCAGGTG 2327  
QY 201 AlaHisLeuHisAlaProThrGlySerGlySerThrLysValProAlaAlaTyrAla 220

DB 2328 GCTCACCTCCATGCTCCACAGGCGGCAAAAGCACCAAGGTCGCCGCTGCATATCA 2387  
QY 221 AlaGlnGlyTyrLysValLeuValLeuAsnProSerValAlaAlaThrMetGlyPheGly 240  
DB 2388 GCTCAGGGCTATAAGGTGCTAGTACTCAACCCCTCTGTGTGTGCAACACTGGGCTTGTG 2447  
QY 241 AlaTyrMetSerLysAlaHisGlyIleAspProAsnIleArgThrGlyValArgThrIle 260  
DB 2448 GCTTACATGTCCAAGGCTCATGGATCGATCTTAACATCAGACCGGGGTGAGAACAATT 2507  
QY 261 ThrThrGlySerProIleThrTyrSerThrTyrGlyLysPheLeuAlaAspGlyGlyCys 280  
DB 2508 ACCACTGGACGCCCATCAGTATCTCCACTACGGCAAGTTCTTGGCGAGCGGGGTGC 2567  
QY 281 SerGlyGlyAlaTyrAspIleIleCysAspGluCysHisSerThrAspAlaThrSer 300  
DB 2568 TCGGGGGCGCTTATGACATANTATTTGTGACGAGTGCCACTCCACGGATGCCACATCC 2627  
QY 301 IleLeuGlyIleGlyThrValLeuAspGlnAlaGluThrAlaGlyAlaArgLeuThrVal 320  
DB 2628 ATCTTGGGATCGGCACCTGCTTGAACAAGCAGAGACTGGCGGGCGAGACTGGTTGTG 2687  
QY 321 LeuAlaThrAlaThrProGlySerValThrValProHisProAsnIleGluGluVal 340  
DB 2688 CTGCGCACCGCCACCTCTCGGGCTCCGTCTCACTGTGCCCATCCCAACATCGAGAGGTT 2747  
QY 341 AlaLeuSerThrThrGlyGluIleProPheTyrGlyLysAlaIleProLeuGluAlaIle 360  
DB 2748 GCTCTGTCCACCACCGAGAGATCCCTTTTACGGCAAGGCTATCCCTCGAGTAATC 2807  
QY 361 LysGlyGlyArgHisLeuIlePheCysHisSerLysLysCysAspGluLeuAlaAla 380  
DB 2808 AAGGGGGGAGACATCTCATCTTCTGTCATTCAAGAAGAAGTGCAGCAACTCGCGCA 2867  
QY 381 LysLeuValAlaLeuGlyValAsnAlaValAlaTyrTyrArgGlyLeuAspValSerVal 400  
DB 2868 AAGTGTGTCGATTTGGGCATCAATCGCGTGTCTACTACCGCGTCTTTGACGTGTCGTC 2927  
QY 401 IleProThrSerGlyAspValValValValAlaThrAspAlaLeuMetThrGlyPheThr 420  
DB 2928 ATCCGACACGCGCGATGTTGTCGTGTGCAACCATGATGCCCTCATGACCGCTATACC 2987  
QY 421 GlyAspPheAspSerValIleAspCysAsnThrCysValThrGlnThrValAspPheSer 440  
DB 2988 GGCACCTTCGACTCGGTGATAGACTGCAATACGTGTGTACCCAGACAGTGCATTTTACG 3047  
QY 441 LeuAspProThrPheThrIleGluThrIleThrLeuProGlnAspAlaValSerArgThr 460  
DB 3048 CTTGACCTTACTCTTACCATTTGAGACAATCAGCTCCCGCAGGATGCTGTCTCCGCACT 3107  
QY 461 GlnArgArgGlyArgThrGlyValGlyValGlyValGlyValGlyValGlyValGlyVal 480  
DB 3108 CAAGTGGGGGAGAGACTGGCAGGGGAAGCCAGGCATCTACAGATTGTGGACCGGGG 3167  
QY 481 GluArgProSerGlyMetPheAspSerSerValLeuCysGluCysTyrAspAlaGlyCys 500  
DB 3168 GAGCGCCCTCCGCGATGTTTCGACTGCTGCTCTCTGTGAGTGTATGACGAGGCTGT 3227  
QY 501 AlaTrpTyrGluLeuThrProAlaGluThrThrValArgLeuArgAlaTyrMetAsnThr 520  
DB 3228 GCTTGGTATGAGCTCAGCGCCCGCAGACTACAGTAGGCTACGAGGTACGAGGTACGAACACC 3287  
QY 521 ProGlyLeuProValCysGlnAspHisLeuGluPheTrpGluGlyValPheThrGlyLeu 540  
DB 3288 CCGGGGCTTCCCGGTGCCAGGACCATCTTGAAATTTGGGAGGGCGCTTTTACAGGCTC 3347  
QY 541 ThrHisIleAspAlaHisPheLeuSerGlnThrLysGlnSerGlyGluAsnLeuProTyr 560  
DB 3348 ACTCATATAGTGCCTTCTTATCCAGACAAGCAGAGTGGGGAGAACCTTCTTAC 3407  
QY 561 LeuValAlaTyrGlnAlaThrValCysAlaArgAlaGlnAlaProProSerTrpAsp 580

Db 3408 CTGTAGCGTACCAGCCACCGTGTGCGCTAGGCTCAAGCCCTCCCTCCCATCGTGGGAC 3467  
 Qy 581 GlnMetTrpLysCysLeuIleAsgLeuLysProThrLeuHisGlyProThrProLeuLeu 600  
 Db 3468 CAGATGTGGAGAGTGTGGCTCAGCCACCTCCATGGCCCAACACCCCTGCTA 3527  
 Qy 601 TTTArgLeuGlyAlaValGlnAsgLeuValThrLeuThrHisProValThrLysTyrIle 620  
 Db 3528 TACAGACTGGCGCTGTTCAAGATGAATCACCCTGAGCGCACCCAGTCCACCAATACATC 3587  
 Qy 621 MetThrCysMetSerAlaAspLeuGluValValThrSerThrTrpValLeuValGlyGly 640  
 Db 3588 ATGACATGCATGTGCGCGCACCTGGAGTGTGTCACGAGCACCTGGGTGCTCGTGGCGGC 3647  
 Qy 641 ValLeuAlaAlaLeuAlaTyrCysLeuSerThrGlyCysValValIleValGlyArg 660  
 Db 3648 GTCTGCTGCTTTGGCGCGTATGTCTGTCAACAGCTGCGTGGTCAATAGTGGGAGG 3707  
 Qy 661 IleValLeuSerGlyLysProAlaIleIleProAspArgGluValLeuTyrArgGluPhe 680  
 Db 3708 GTGCTGTGTCCGGAAGCGGCAATCATACCTGACAGGGAAGTCTCTACCGAGAGTTC 3767  
 Qy 681 AspGluMetGluGluCys 686  
 Db 3768 GATGAGATGGAAGAGTGC 3785

## RESULT 7

AAAN90336  
 ID AAAN90336 standard; DNA; 7310 BP.

XX AC AAAN90336;

XX 25-MAR-2003 (revised)

DT 19-JUL-2001 (revised)

DT 01-NOV-1989 (first entry)

XX Composite hepatitis C virus (HCV) cDNA.

XX Hepatitis C virus; cDNA; clone 15e; clone k9-1; probe; vaccine; ds.

XX Pan troglodytes.

XX GB2212511-A.

XX 26-JUL-1989.

PF 18-NOV-1988; 88GB-00027024.

XX 18-NOV-1987; 87US-00122714.

PR 30-DEC-1987; 87US-00139886.

PR 26-FEB-1988; 88US-00161072.

PR 26-OCT-1988; 88US-00263584.

XX (CHIR ) CHIRON CORP.

PA Houghton M, Choo QL, Kuo G;

XX WPI; 1989-215054/30.

DR P-PSDB; AAP90288.

XX Hepatitis C virus gene - used for prodn. of polynucleotide probes  
 PT polypeptide(s) and antibodies for diagnosis, prevention and treatment of  
 PT infection.

XX Disclosure; Fig 47; 30pp; English.

XX The sequence shows a composite hepatitis C virus (HCV) cDNA, derived by  
 CC aligning clones k9-1 through 15e in 5'-3' direction. The cDNA encodes  
 CC antigens which react with antibodies in patients with non-A non-B  
 CC hepatitis (NANBH). The cDNA can be used to design probes, or to  
 CC synthesise polypeptides, which are used to diagnose HCV-induced NANBH, to  
 CC raise antibodies for immunoassay or treatment, or to produce vaccines.  
 CC See also AAP90288, and AAAN90303-35. (N.B. This record was resubmitted to

CC correct errors in the sequence.) (Updated on 25-MAR-2003 to correct PR  
 CC field.)  
 XX  
 SQ Sequence 7310 BP; 1495 A; 2218 C; 2058 G; 1539 T; 0 U; 0 Other;  
 Alignment Scores:  
 Pred. No.: 1-88e-241 Length: 7310  
 Score: 3574.00 Matches: 672  
 Percent Similarity: 99.42% Conservative: 10  
 Best Local Similarity: 97.96% Mismatches: 4  
 Query Match: 98.78% Indels: 0  
 DB: 1 Gaps: 0  
 US-09-930-591-2 (1-686) x AAAN90336 (1-7310)  
 Qy 1 MetAlaProIleThrAlaTyrAlaGlnGlnThrArgGlyLeuLeuGlyCysIleIleThr 20  
 Db 1728 CTGGCCCATCACGGCGTACGCCAGACAGCAAGGGGCTCTTAGGGTGCAATACCC 1787  
 Qy 21 SerLeuThrGlyVarAspLysAenGlnValGluGlyValGlnIleValSerThrAla 40  
 Db 1788 AGCTTAACTGGCGCGGACAAACCAAGTGGAGGTGAGGTCCAGATTGTCTCACTGCT 1847  
 Qy 41 AlaGlnThrPheLeuAlaThrCysIleAenGlyValCysTrpThrValTyrHisGlyAla 60  
 Db 1848 GCCAAACCTTCTGGCAACGTGCATCAATGGGTGTGTGGACTGTCTACCAAGGGGCC 1907  
 Qy 61 GlyThrArgThrIleAlaSerProLysGlyProValIleGlnMetTyrThrAsnValAsp 80  
 Db 1908 GGAACGAGACCATCGCGTCACCAAGGCTCTGTCAATCCAGATGTATACCAATGTAGAC 1967  
 Qy 81 GlnAspLeuValGlyTyrProAlaProGlnGlyAlaArgSerLeuThrProCysThrCys 100  
 Db 1968 CAAGACCTTGTGGGTGGCGCGCTCGCAAGTAGACCGCTCATTTGACACCTTGCACTTGC 2027  
 Qy 101 GlySerSerAspLeuTyrLeuValThrArgHisAlaAspValIleProValArgArgArg 120  
 Db 2028 GGTCTCTCGGACCTTTACCTGTGTACAGGACGCGCATGTCTCCGTCCGCGCGCGG 2087  
 Qy 121 GlyAspGlyArgGlySerLeuLeuSerProArgProIleSerTyrIleLysGlySerSer 140  
 Db 2088 GGTGATAGCAGGGGCGAGCTGTGTGCGCCCGGCCCATTTCTTACTTTGAAAGCTCTCTCG 2147  
 Qy 141 GlyGlyProLeuLeuCysProAlaGlyHisAlaValGlyIlePheArgAlaAlaValCys 160  
 Db 2148 GGGGTCCGCTGTGTGCGCGCGGCGCACGCGTGGGCATATTTAGGCGCGCGGTGTC 2207  
 Qy 161 ThrArgGlyValAlaLysAlaValAspPheIleProValGluSerLeuGluThrThrMet 180  
 Db 2208 ACCGTGGAGTGGCTAAGGCGGTGGACTTTATCCCTGTGGAGAACCTAGACACCAACATG 2267  
 Qy 181 ArgSerProValPheSerAspAsnSerSerProAlaValProGlnSerTyrGlnVal 200  
 Db 2268 AGGTCCCGGTGTTCACGGATAACTCTCTCCACAGTAGTGTCCCGCAGAGTTCACAGGTG 2327  
 Qy 201 AlaHisLeuHisAlaProThrGlySerGlyLysSerThrLysValProAlaAlaTyrAla 220  
 Db 2328 GTTCACTCCATGCTCCACAGCGCGGCAAAAGCACCAAGGTCCCGCTGCATATGCA 2387  
 Qy 221 AlaGlnGlyTyrLysValLeuValLeuAsnProSerValAlaAlaThrMetGlyPheGly 240  
 Db 2388 GCTCAGGGCTATAAGGTGTAGTACTCAACCCCTCTGTGTGTGTCGAACACTGGGCTTTGCT 2447  
 Qy 241 AlaTyrMetSerLysAlaHisGlyIleAspProAsnIleArgThrGlyValArgThrIle 260  
 Db 2448 GCTTACATGTCGAAGGCTCATGGATCGATCTTAACATCAGGACCGGGGTGAGAACAT 2507  
 Qy 261 ThrThrGlySerProIleThrTyrSerThrTyrGlyLysPheLeuAlaAspGlyGlyCys 280  
 Db 2508 ACCATGGCAGCGCCCATCAGTACTCCACTACGCAAGTTCCTTGGCGCGCGGTGTC 2567  
 Qy 281 SerGlyGlyAlaTyrAspIleIleleCysAspGluCysHisSerThrAspAlaThrSer 300

```
Db 2568 TCGGGGGCGCTTATGACATATAATTGTGACGAGTCCACTCCACGATGCCATCC 2627
Qy 301 lleLeuGlyileGlyThrValLeuAspGlnAlaGluThrAlaGlyValAArgLeuThrVal 320
Db 2628 ATCTTGGGCATCGGCACCTGCTTGTACCAAGCAGAGACTGCGGGGCGAGACTGGTTGTG 2687
Qy 321 LeuAlaThrAlaThrProProGlySerValThrValProHisProAsnIleGluGluVal 340
Db 2688 CTCGCCACCGCACCCCTCCGGGCTCCGCTACTGTGCCCCATCCCAACATCGAGAGGTT 2747
Qy 341 AlalaLeuSerThrThrGlyGluLeuProPheThrGlyLysAlaIleProLeuGluAlaIle 360
Db 2748 GCTCTGTCCACCGCAGAGATCCCTTTTACGCAAGGCTATCCCTCGAAGTAATC 2807
Qy 361 LysGlyGlyArgHisLeuIlePheCysHisSertyLysLysCysAspGluLeuAlaAla 380
Db 2808 AAGGGGGGAGACATCTCATCTTCTGTCATTCAAGAAGAGTGCAGCAATCGCGCA 2867
Qy 381 LysLeuValAlaLeuGlyValAsnAlaValAlaIleValThrValArgGlyLeuAspValSerVal 400
Db 2868 AAGCTGTGTGCGATTTGGGATCAATCGCTGGCTACTACCGGGTCTTTGACGTGTCCGTC 2927
Qy 401 IleProThrSerGlyAspValValValValAlaThrAspAlaLeuMetThrGlyPheThr 420
Db 2928 ATCCCGACCGCGGGATGTTGTGCTGTGGCAACCGATGCCCTCATGCCGGCTATACC 2987
Qy 421 GlyAspPheAspSerValIleAspCysAsnThrCysValThrGlnThrValAspPheSer 440
Db 2988 GCGGACTTCGACTCGGTGATAGACTGCAATACGTGTGTACCCAGACAGTGCATTTACG 3047
Qy 441 LeuAspProThrPheThrIleGluThrIleThrLeuProGlnAspAlaValSerArgThr 460
Db 3048 CTTGACCCCTACCTTACCATTGAGACAATCAGCTCCGCCAGGATGCTGTCTCCGCACT 3107
Qy 461 GlnArgArgGlyArgThrGlyArgGlyLysProGlyIleThrArgPheValAlaProGly 480
Db 3108 CAACGTCCGGGCGAGACTGGCAGGGGAGCCAGGCATCTACAGATTTGTGGCACCGGG 3167
Qy 481 GluArgProSerGlyMetPheAspSerSerValLeuCysGluCysTyrAspAlaGlyCys 500
Db 3168 GAGCGCCCTCCGGCATGTTGCACTCGTCGCTCTCTGTGAGTGTATGACGAGGCTGT 3227
Qy 501 AlaTrpTyrGluLeuThrProAlaGluThrThrValArgLeuArgAlaTyrMetAsnThr 520
Db 3228 GCTTGGTATGACTCAGCCGCCCGCAGACTACAGTTAGGCTACGAGCGTACATGAACACC 3287
Qy 521 ProGlyLeuProValCysGlnAspHisLeuGluPheTrpGluGlyValPheThrGlyLeu 540
Db 3288 CCGGGGCTCCCGTGTCCAGGACCATCTTGAATTTTGGGAGGGCGCTTTTACAGGCTC 3347
Qy 541 ThrHisIleAspAlaHisPheLeuSerGlnThrLysGlnSerGlyGluAsnLeuProTyr 560
Db 3348 ACTCATATAGATGCCACTTTCTATCCACACAAGACAGAGTGGGGAGAACCTTCTTAC 3407
Qy 561 LeuValAlaTyrGlnAlaThrValCysAlaArgAlaGlnAlaProProSerTrpAsp 580
Db 3408 CTGGTAGCGTACCAAGCCACCGTGTGCGTAGGCTCAAGGCCCTCCCCCATCGTGGAC 3467
Qy 581 GlnMetTrpLysCysLeuIleArgLeuLysProThrLeuHisGlyProThrProLeuLeu 600
Db 3468 CAGATGTGGAAGTGTGTGATTCCGCTCAAGCCACCTCCATGGGCCAACACCCCTGTGA 3527
Qy 601 TyrArgLeuGlyValAlaValGlnAsnGluValThrLeuThrHisProValThrLysTyrIle 620
Db 3528 TACAGACTGGGGCGCTTTCAGAAATGAAATCACCTTGACGCCACCCAGTCCCAATACATC 3587
Qy 621 MetThrCysMetSerAlaAspLeuGluValValThrSerThrTrpValLeuValGlyGly 640
Db 3588 ATGATCATGATGTCCGGGACCTGGAGGTCTGTCACGAGCACCTGGGTGCTCGTTGGCGGC 3647
Qy 641 ValLeuAlaAlaLeuAlaTyrCysLeuSerThrGlyCysValValIleValGlyArg 660
Db 3648 GTCCCTGGCTGCTTTGGCGCGGTATTTCCTGTCAACAGGCTGCGTGTGATAGTGGCAGG 3707
```

```
Qy 661 lleValLeuSerGlyLysProAlaIleIleProAspArgGluValLeuTyrArgGluPhe 680
Db 3708 GTCTCTGTTCGCGGAAGCCGCAATCATATACCTCACAGGAAGTCTCTTACCGAGATTC 3767
Qy 681 AspGluMetGluGluCys 686
Db 3768 GATGAGATGGAAGAGATGC 3785
RESULT 8
ID AAQ98221 standard; cDNA to mRNA; 7310 BP.
XX AAQ98221;
AC AAQ98221;
DT 25-MAR-2003 (revised)
DT 15-AUG-1996 (first entry)
XX Hepatitis C virus clone genome.
DE Hepatitis C virus; HCV; antigen; detection; diagnosis; vaccine;
KW antibodies; immunoprophylaxis; sera; serum; ds.
XX Hepatitis C virus.
OS US5443965-A.
PN 22-AUG-1995.
XX 05-APR-1991; 91US-00681703.
XX 06-APR-1990; 90US-00505611.
PR 09-OCT-1990; 90US-00594854.
XX (GENE-) GENELABS INC.
XX Mosekli R, Reyes GR, Kim JP;
PI WPI; 1995-302120/39.
DR New nucleic acids encoding hepatitis C virus antigens - used to develop
PT prods. for detection of HCV-infected sera and prodn. of vaccines and anti
PT -HCV antibodies.
PS Example 4; Fig 11; 71pp; English.
XX Hepatitis C virus (HCV) antigens can be used for detecting HCV infected
CC sera and individuals infected with HCV. They can also be used in an anti-
CC HCV vaccine or for the production of anti-HCV antibodies which can be
CC used for passive immunoprophylaxis. The antigens consistently identify
CC more HCV positive serum samples with a high degree of specificity. See
CC AAQ98202-14 and AAR81939-51. (Updated on 25-MAR-2003 to correct PF
CC field.) (Updated on 25-MAR-2003 to correct PR field.)
XX
SQ Sequence 7310 BP; 1494 A; 2217 C; 2060 G; 1539 T; 0 U; 0 Other;
Alignment Scores:
Pred. No.: 1.88e-241 Length: 7310
Score: 3574.00 Matches: 672
Percent Similarity: 99.42% Conservative: 10
Best Local Similarity: 97.96% Mismatches: 4
Query Match: 98.78% Indels: 0
DB: Gaps: 0
US-09-930-591-2 (1-686) x AAQ98221 (1-7310)
Qy 11 MetAlaProIleThrAlaTyrAlaGlnGlnThrArgGlyLeuLeuGlyCysIleIleThr 20
Db 1728 CTGGCGCCCATCGCGGTACGCCAGCAGACAAGGGCCCTCTTAGGGTGCAATATCACC 1787
Qy 21 SerLeuThrGlyArgAspLysAsnGlnValGluGlyGluValGlnIleValSerThrAla 40
Db 1788 AGCCTTAAGTCCCGGACAAAAAACAAGTGGAGGGTGAGGTCAGATTGTGTCAACTGCT 1847
```

```
QY 41 AlaGlnThrPheLeuAlaThrCysIleAsnGlyValCysTrpThrValTyrHisGlyVala 60
Db 1848 GCCCAAACTTCCTGGCAACGTCATCAATGGGGTGTGCTGGACTGTCTACCGGGGCC 1907

QY 61 GlyThrArgThrIleAlaSerProLysGlyProValIleGlnMetTyrThrAsnValAsp 80
Db 1908 GGAACGAGGACCATCGCGTCACCCAGGGTCCTGTCTATCCAGATGTATACCAATGTAGAC 1967

QY 81 GluAspLeuValGlyTyrProAlaProGlnGlyAlaArgSerLeuThrProCysThrCys 100
Db 1968 CAAGACCTTGTGGGCTGGCCGCTCCGCAAGGTAGCGCTCATGACACCTCGACTTGC 2027

QY 101 GlySerSerAspLeuTyrLeuValThrArgHisAlaAspValIleProValArgArgArg 120
Db 2028 GGCTCCTCGGACCTTACTTGTGTACGAGGACGCCGATGTCAATCCCGTGGCGGGCGG 2087

QY 121 GlyAspGlyArgGlySerLeuLeuSerProArgProIleSerTyrLeuLysGlySerSer 140
Db 2088 GGTGATAGCAGGGGACGCTGTGTGCCCCCGGCCCAATTCCTACTTGAAGGCTCCTCG 2147

QY 141 GlyGlyProLeuLeuCysProAlaGlyHisAlaValGlyIlePheArgAlaAlaValCys 160
Db 2148 GGGGCTCCGCTGTGTGCCCGGGGGGACGCCGTGGGCATATTAGGGCCGGGGTGTGC 2207

QY 161 ThrArgGlyValAlaLysAlaValAspPheIleProValGluSerLeuGluThrThrMet 180
Db 2208 ACCGTGGAGTGGCTAAGCGGTGGACTTTATCCCTGTGGAGAACCTTAGAGACACCATG 2267

QY 181 ArgSerProValPheSerAspAsnSerSerProProAlaValProGlnSerTyrGlnVal 200
Db 2268 AGGTCCCGGTGTTCACGGATACTCTCTCCACCATAGTAGTGGCCACAGAGCTTCCAGGTG 2327

QY 201 AlaHisLeuHisAlaProThrGlySerGlyLysSerThrLysValProAlaAlaTyrAla 220
Db 2328 GCTACCTCCTCATGCTCCACAGGAGGGCAAGACCAAGGTCCCGGCTGCATATGCA 2387

QY 221 AlaGlnGlyTyrLysValLeuValLeuAsnProSerValAlaAlaThrMetGlyPheGly 240
Db 2388 GCTCAGGGCTATAAGGTGTAGTACTCAACCCCTCTGTGTCTGCACACTGGGCTTTGGT 2447

QY 241 AlaTyrMetSerLysAlaHisGlyIleAspProAsnIleArgThrGlyValArgThrIle 260
Db 2448 GCTTACATGTCCAGGCTCATGGATCGATCCTAAACATCAGGACCGGGGTGAGAACAAAT 2507

QY 261 ThrThrGlySerProIleThrTyrSerThrTyrGlyLysPheLeuAlaAspGlyGlyCys 280
Db 2508 AACACTGGCAGCCCATCAGTACTCCACTACGGCAAGTTCCTTGCACGCGGGGTGC 2567

QY 281 SerGlyGlyAlaTyrAspIleIleCysAspGluCysHisSerThrAspAlaThrSer 300
Db 2568 TCGGGGGCGCTTATGACATAATAATTTGTGACGAGTGGCCACTCCACGGATGCCATCC 2627

QY 301 IleLeuGlyIleGlyThrValLeuAspGlnAlaGluThrAlaGlyAlaArgLeuThrVal 320
Db 2628 ATCTTGGGCATCGGCACTGCTTGAACCAAGCAGAGACTCGGGGGCGAGACTGGTTGTG 2687

QY 321 LeuAlaThrAlaThrProGlySerValThrValProHisProAsnIleGluGluVal 340
Db 2688 CTGCGCACCCGCCACCTCCGGGCTCCGTCACTGTGCCCATCCCAACATCGAGGAGGTT 2747

QY 341 AlaLeuSerThrThrGlyGluIleProPheTyrGlyLysAlaIleProLeuGluAlaIle 360
Db 2748 GCTCTGTCCACCGGAGAGATCCCTTTTACGGCAAGGCTATCCCTCTCGAAGTAATC 2807

QY 361 LysGlyGlyArgHisLeuIlePheCysHisSerLysLysLysCysAspGluLeuAlaAla 380
Db 2808 AAGGGGGGAGACATCTCATCTTCTGTCAATTCAAAGAGAAGTGCAGCAACTCGCGCA 2867

QY 381 LysLeuValAlaLeuGlyValAsnAlaValAlaTyrTyrArgGlyLeuAspValSerVal 400
Db 2868 AAGCTGTGTCANTGGGCATCAATGCGTGGCTTACTTACCGCGGTCTTGACGTGTCGTC 2927
```

```
QY 401 IleProThrSerGlyAspValValValAlaThrAspAlaLeuMetThrGlyPheThr 420
Db 2928 ATCCCGACAGGGCGCATGTGTGTCGTGGCAACGATCCCTCATGACCGGCTATACC 2987

QY 421 GlyAspPheAspSerValIleAspCysAsnThrCysValThrGlnThrValAspPheSer 440
Db 2988 GCGACTTCGACTCGGTGATAGACTCAATACGTGTGTGCCAGACAGTCGATTTTCAGC 3047

QY 441 LeuAspProThrPheThrIleGluThrIleThrLeuProGlnAspAlaValSerArgThr 460
Db 3048 CTTGACCCCTTACCTTACCATTTGAGACAATCACGCTCCCCAGGATGCTGTCTCCGACT 3107

QY 461 GlnArgArgGlyArgThrGlyArgGlyLysProGlyIleTyrArgPheValAlaProGly 480
Db 3108 CAACTGCGGGGAGGAGCTGGCAGGGGGAAGCAGGCACTACAGATTTGTGGCACCGGGG 3167

QY 481 GluArgProSerGlyMetPheAspSerSerValLeuCysGluCysTyrAspAlaGlyCys 500
Db 3168 GAGCGCCCTCCGGCATGTTTCGACTCGTCCGCTCTGTGAGTGTCTATGACGAGGCTGT 3227

QY 501 AlaTyrTyrGluLeuThrProAlaGluThrThrValArgLeuArgAlaTyrMetAsnThr 520
Db 3228 GCTTGTGTATGAGTCTACGCCCGCCGAGACTACAGTTAGGCTACGAGCGTACATGAACAC 3287

QY 521 ProGlyLeuProValCysGlnAspHisLeuGluPheTyrGluGlyValPheThrGlyLeu 540
Db 3288 CCGGGGCTTCCGCTGTGCCAGGACCATCTTGAATTTTGGGAGGGCTCTTTACAGGCCCTC 3347

QY 541 ThrHisIleAspAlaHisPheLeuSerGlnThrLysGlnSerGlyGluAsnLeuProTyr 560
Db 3348 ACTCATATAGATGCCCACTTCTTATCCAGACAAAGCAGAGTGGGAGAACCTTCTCTTAC 3407

QY 561 LeuValAlaTyrGlnAlaThrValCysAlaArgAlaGlnAlaProProProSerTyrAsp 580
Db 3408 CTGGTAGCTACCAAGCCACCGTGTGCGCTAGGGCTCAAGCCCTCCCATCTCGTGGGAC 3467

QY 581 GlnMetTyrLysCysLeuIleArgLeuLysProThrLeuHisGlyProThrProLeuLeu 600
Db 3468 CAGATGTGGAAGTGTGTTGATTGCGCTCAAGCCACCCCTCCATGGGCCAACACCCCTGCTA 3527

QY 601 TyrArgLeuGlyAlaValGlnAsnGluValThrLeuThrHisProValThrLysTyrIle 620
Db 3528 TACAGACTCGGGCGCTGTTTCAGAAATGAATCAACCTGACGCCACCCAGTCCCAAAATACATC 3587

QY 621 MetThrCysMetSerAlaAspLeuGluValValThrSerThrTyrValLeuValGlyGly 640
Db 3588 ATGACATGTCATGTCCGCCCGACCTGGAGGTCTGTACAGAGCACCTGGGTGCTCGTTGGCGGC 3647

QY 641 ValLeuAlaAlaLeuAlaTyrCysLeuSerThrGlyCysValValIleValGlyArg 660
Db 3648 GTCTCGCTGCTTTGGCCGCGTATTGCTGTCAACAGGCTGCGTGTGTATAGTGGGCAGG 3707

QY 661 IleValLeuSerGlyLysProAlaIleIleProAspArgGluValLeuTyrArgGluPhe 680
Db 3708 GTCTGTCTGTCCGGGAAGCGGCAATCATACCTGACAGGGAAGTCTCTTACCGAGAGTTC 3767

QY 681 AspGluMetGluGluCys 686
Db 3768 GATGAGATGGAAGATGTC 3785

RESULT 9
AAA75296
ID AAA75296 standard; cDNA; 8316 BP.
XX
AC AAA75296;
XX
DT 15-JAN-2001 (first entry)
XX
DE cDNA sequence compiled Hepatitis C virus cDNA clones.
XX
KW Hepatitis C virus; HCV; antisense polynucleotide; polypeptide;
XX
XX viral infectivity; viral replication; ds.
```

```

OS Hepatitis C virus.
XX Key Location/Qualifiers
FH 1. .8316
FT CDS /*tag= a
FT /*note= "partial sequence; no termination codon given"
XX
XX EPI034785-A2.
XX 13-SEP-2000.
XX
XX 16-MAR-1990; 2000EP-00109602.
XX
XX 17-MAR-1989; 89US-00325338.
XX 20-APR-1989; 89US-00341334.
XX 18-MAY-1989; 89US-00355002.
XX 16-MAR-1990; 90EP-00302866.
XX
XX (CHIR ) CHIRON CORP.
XX
XX Houghton M, Choo Q, Ruo G;
XX
XX WPI; 2000-566891/53.
XX P-PSDB; AAB18540.
XX
XX Novel composition comprising a hepatitis C virus antisense polynucleotide
XX PT which is complementary to or corresponds to a sense strand of the virus
XX PT genome, and selectively hybridizes to it.
XX
XX Example; Fig 16; 75pp; English.
XX
XX The specification describes a pharmaceutical composition which comprises
XX CC a hepatitis C virus (HCV) antisense polynucleotide. The HCV is
XX CC characterized by a positive stranded RNA genome which has 40% homology at
XX CC the polypeptide level to a HCV polypeptide. The antisense polynucleotide
XX CC binds to cellular polynucleotides which enhance and/or are required for
XX CC viral infectivity, replicative ability or chronicity. The antisense
XX CC polynucleotides may also be designed to bind with high specificity, to be
XX CC of increased stability, to be stable and to have low toxicity. The
XX CC composition also comprises an agent which causes viral RNA to be
XX CC inactive. The composition is used for preventing HCV replication in a
XX CC system. The present sequence represents a novel HCV cDNA sequence, which
XX CC is used in the course of the invention
XX
XX SQ Sequence 8316 BP; 1671 A; 2529 C; 2345 G; 1771 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 2.2e-241 Length: 8316
Score: 3574.00 Matches: 672
Percent Similarity: 99.42% Conservative: 10
Best Local Similarity: 97.96% Mismatches: 4
Query Match: 98.78% Indels: 0
DB: 3 Gaps: 0

US-09-930-591-2 (1-686) x AAA75296 (1-8316)
QY 1 MetAlaProIleThrAlaTyrAlaGlnGlnThrArgGlyLeuLeuGlyCysIleIleThr 20
DB 2734 CTGGCGCCCATCAGCGGTACGCCAGCAGCAAGGGCCCTCTTAGGGTGATTAATCACC 2793
QY 21 SerLeuThrGlyArgAspLysAsnGlnValGluGlyGluValGlnIleValSerAla 40
DB 2794 AGCCTAACTGCCCGGGACAAAACCAAGTGAGGGTGAGGTCCAGATTGTGTCAACTGCT 2853
QY 41 AlaGlnThrPheLeuAlaThrCysIleAsnGlyValCysTrpThrValTyrHisGlyAla 60
DB 2854 GCCCAAACTTCTTGGCAACGTGATCAATGGGGTGTCTGAGACTGTCTACACGGGGCC 2913
QY 61 GlyThrArgThrIleAlaSerProLysGlyProValIleGlnMetTyrThrAsnValAsp 80
DB 2914 GGAACAGGACCATTCGGCTACCCAGGGTCTCTGATCCAGAGTATACCAATGTAGAC 2973
QY 81 GlnAspLeuValGlyTrpProAlaProGlnGlyAlaArgSerLeuThrProCysThrCys 100

```

```

Db 2974 CAAGACCTTGTGGCTGGCCCGCTCCGCAAGGTAGCGCTCATTTGACACCTTGACCTTGC 3033
QY 101 GlySerSerAspLeuTyrLeuValThrArgHisAlaAspValIleProValArgArg 120
DB 3034 GGCTCTCGGACCTTACCTGGTTCAGGAGCAGCCGATGTCTATCCGTGGCGGGCGG 3093
QY 121 GlyAspGlyArgGlySerLeuLeuSerProArgProIleSerTyrLeuLysGlySerSer 140
DB 3094 GGTGATAGCAGGGGCGAGCTGCTCGCCCGGCCATTTCTACTTTGAAGGCTCTCTCG 3153
QY 141 GlyGlyProLeuLeuCysProAlaGlyHisAlaValGlyIlePheArgAlaAlaValCys 160
DB 3154 GGGGCTCGCTGTGTGCTCCCGGGGACGCCGTGGGCATATTTAGGGCGCGGTGTC 3213
QY 161 ThrArgGlyValAlaLysAlaValAspPheIleProValGluSerLeuGluThrThrMet 180
DB 3214 ACCGTGAGTGGCTAAGCGGTGACTTTATCTCTGTGGAGAACCTTAGACCAACCATG 3273
QY 181 ArgSerProValPheSerAspAsnSerSerProProAlaValProGlnSerTyrGlnVal 200
DB 3274 AGTCCCGGCTTTCACGGATAACTCTCTCCACAGTAGTGGCCCGACAGCTTCAGGTG 3333
QY 201 AlaHisLeuHisAlaProThrGlySerGlyLysSerThrLysValProAlaAlaTyrAla 220
DB 3334 GCTCACCTCCATGCTCCACAGGCGGCAAGAACCAAGAGTCCCGGCTGCAATATGCA 3393
QY 221 AlaGlnGlyTyrLysValLeuValLeuAsnProSerValAlaAlaThrMetGlyPheGly 240
DB 3394 GCTCAGGGCTATAGGTGTAGTACTCAACCCCTCTGTGTGTCGAACACTGGGCTTGGT 3453
QY 241 AlaTyrMetSerLysAlaHisGlyIleAspProAsnIleArgThrGlyValArgThrIle 260
DB 3454 GCTTACATGTCCAAGGCTCATGGGATCGATCTTAACATCAGGACCGGGTGAGAACATT 3513
QY 261 ThrThrGlySerProIleThrTyrSerThrTyrGlyLysPheLeuAlaAspGlyGlyCys 280
DB 3514 ACCACTGGCAGCCCATCAGCTACTCCACCTACGCAAGTTCTTTCGCGAGCGGGGTC 3573
QY 281 SerGlyGlyAlaTyrAspIleIleCysAspGluCysHisSerThrAspAlaThrSer 300
DB 3574 TCGGGGGCGCTTATGACATTAATTTGTGACAGTGCCTCCAGGATGCCATCCAGATCC 3633
QY 301 IleLeuGlyIleGlyThrValLeuAspGlnAlaGluThrAlaGlyAlaArgLeuThrVal 320
DB 3634 ATCTTGGGCTCGGCACTGCTCTTGACCAAGCAGAGACTGCGGGGGCGAGACTGTTGTG 3693
QY 321 LeuAlaThrAlaThrProProGlySerValThrValProHisProAsnIleGluGluVal 340
DB 3694 CTGCGCACCGCCACCCCTCGGGGCTCGGTCACTGTGCCCCCATCCCAACATCGAGAGGTT 3753
QY 341 AlaLeuSerThrThrGlyGluIleProPheTyrGlyLysAlaIleProLeuGluAlaIle 360
DB 3754 GCTCTGCCACCCCGGAGAGATCCCTTTTACGGCAAGGCTATCCCCCTCGAAGTAATC 3813
QY 361 LysGlyGlyArgHisLeuIlePheCysHisSerLysLysCysAspGluLeuAlaAla 380
DB 3814 AAGGGGGGAGACATCTCATCTCTGTCTTCAAAAGAAAGTGCAGCAACTCGCCGCA 3873
QY 381 LysLeuValAlaLeuGlyValAsnAlaValAlaTyrTyrArgGlyLeuAspValSerVal 400
DB 3874 AAGTGTGTCATTTGGGCATCAATCGGTGGCTTACTACCGCGGTCTTACGTCTCGTGC 3933
QY 401 IleProThrSerGlyAspValValValAlaThrAspAlaLeuMetThrGlyPheThr 420
DB 3934 ATCCCGACCGCGCGATGTTGCTGTCGTCACCAACCATGCTCCATGACCGGCTATACC 3993
QY 421 GlyAspPheAspSerValIleAspCysAsnThrCysValThrGlnThrValAspPheSer 440
DB 3994 GGGCACTTCGACTCGGTGATAGACTGCAATACGTGTGTACCCAGACAGATCGATTCAGC 4053
QY 441 LeuAspProThrPheThrIleGluThrIleThrLeuProGlnAspAlaValSerArgThr 460

```





Qy 101 GlySerSerAspLeuTyrLeuValThrArgHisAlaAspValIleProValArgArg 120  
Db 3643 GGCTCTCGGACCTTTACCTGGTCACGAGGACGCCGATGTCTCCCGTGGCGGG 3702  
Qy 121 GlyAspGlyArgGlySerLeuLeuSerProArgProIleSerTyrLeuGlySerSer 140  
Db 3703 GGTGATACAGGGGAGCGCTGCTGCGCCCGGCCATTTCTCTGAAAGGCTCTCG 3762  
Qy 141 GlyGlyProLeuLeuCysProAlaGlyHisAlaValGlyIlePheArgAlaAlaValCys 160  
Db 3763 GGGGTCGCTGTTGTGTCGCCCGGGGACGCCGTGGGCATATTAGGGCCGCGGTGTC 3822  
Qy 161 ThrArgGlyValAlaLysAlaValAspPheIleProValGluSerLeuGluThrThrMet 180  
Db 3823 ACCGCTGGAGTGGCTAAAGCGGTGACCTTATCCCTGTGGAGAACCTAGAGACAACCATG 3882  
Qy 181 ArgSerProValPheSerAspAsnSerSerProProAlaValProGlnSerTyrGlnVal 200  
Db 3883 AGGTCCCCGGGTGTTCCAGGTAACCTCTCTCCACGAGTAGTGCCCCAGAGCTTCAGGTG 3942  
Qy 201 AlaHisLeuHisAlaProThrGlySerGlyLysSerThrLysValProAlaAlaTyrAla 220  
Db 3943 GCTCACCTCCATGCTCCACAGGCGCGGCAAGACCAAGGTCCCGGCTGCATATGCA 4002  
Qy 221 AlaGlnGlyTyrLysValLeuValLeuAsnProSerValAlaAlaThrMetGlyPheGly 240  
Db 4003 GCTCAGGCTATAAGGTCTAGTACTCAACCCCTCTGTGTGCAACACTGGGCTTTGGT 4062  
Qy 241 AlaTyrMetSerLysAlaHisGlyIleAspProAsnIleArgThrGlyValArgThrIle 260  
Db 4063 GCTTACATGTCACAGGCTCATGGGATCGATCCTTAACATCAGGACCGGGGTGAGAACAT 4122  
Qy 261 ThrThrGlySerProIleThrTyrSerThrTyrGlyLysPheLeuAlaAspGlyGlyCys 280  
Db 4123 ACCACTGGCAGCCCATCATCGTACTCCACCTACGGCAAGTTCCTTGGCAGCGGGGTGC 4182  
Qy 281 SerGlyGlyAlaTyrAspIleIleCysAspGluCysHisSerThrAspAlaThrSer 300  
Db 4183 TCGGGGGCGCTTATGACATAATAATTTGTACAGAGTGCCACTCCAGGATGCCACATCC 4242  
Qy 301 IleLeuGlyIleGlyThrValLeuAspGlnAlaGluThrAlaGlyAlaArgLeuThrVal 320  
Db 4243 ATCTTGGGCATCGGCAGCTGCTCTTGACCAAGCAGAGACTGCGGGGGGAGACTCGTTGTG 4302  
Qy 321 LeuAlaThrAlaThrProGlySerValThrValProHisProAsnIleGluGluVal 340  
Db 4303 CTGCCACCGCCACCCCTCCGGGCTCGCTACTGTGCCCCATCCCAACATCGAGAGGTT 4362  
Qy 341 AlaLeuSerThrThrGlyGluIleProPheTyrGlyLysAlaIleProLeuGluAlaIle 360  
Db 4363 GCTCTGTCCACCCGGAGAGATCCCTTTTACGGCAAGGCTATCCCTCGAAGTAATC 4422  
Qy 361 LysGlyGlyArgHisLeuIlePheCysHisSerLysLysCysAspGluLeuAlaAla 380  
Db 4423 AAGGGGGGAGACATCTCATCTCTGTCTATTCAAAGAGAGAGTGCAGCACTCGCGCA 4482  
Qy 381 LysLeuValAlaLeuGlyValAsnAlaValAlaTyrTyrArgGlyLeuAspValSerVal 400  
Db 4483 AAGCTGGTCGATTTGGGATCAATCGCGTGGCCCTACTACCGCGTCTTGAGTGTCCGTC 4542  
Qy 401 IleProThrSerGlyAspValValValAlaThrAspAlaLeuMetThrGlyPheThr 420  
Db 4543 ATCCCGACCGCGGATGTTGTGCTGTCGCAACCGATGCTCATGACCGGCTATACC 4602  
Qy 421 GlyAspPheAspSerValIleAspCysAsnThrCysValThrGlnThrValAspPheSer 440  
Db 4603 GGCAGCTTCGACTCGGTGATAGATGCAATACGTGTGTACCCAGACAGTGCATTTCCAGC 4662  
Qy 441 LeuAspProThrPheThrIleGluThrIleThrLeuProGlnAspAlaValSerArgThr 460  
Db 4663 CTTGACCTTACCTTTCACATTTGAGACAATACGCTCCCGCAGGATGTGTCTCCCGCACT 4722

Qy 461 GlnArgArgGlyArgThrGlyArgGlyLysProGlyIleTyrArgPheValAlaProGly 480  
Db 4723 CAACGTGGGCGCAGGACTGGCAGGGGGAAGCAGGCATCTACAGATTTGTGGCACCGGG 4782  
Qy 481 GluArgProSerGlyMetPheAspSerSerValLeuCysGluCysTyrAspAlaGlyCys 500  
Db 4783 GAGCGCCCTCCGCGATGTTGACTCGTCCGTCTCTGTGAGTGTATGACGCGAGGCTGT 4842  
Qy 501 AlaTyrTyrGluLeuThrProAlaGluThrThrValArgLeuArgAlaTyrMetAsnThr 520  
Db 4843 GCTTGGTATGAGTCAACGCCCGCCGAGACTACAGTATAGGCTACGAGCGTACATGAACACC 4902  
Qy 521 ProGlyLeuProValCysGlnAspHisLeuGluPheTyrGluGlyValPheThrGlyLeu 540  
Db 4903 CCGGGGCTTCCGCTGTGCCAGGACCATCTTGAATTTTGGGAGGCGCTTTTACAGGCTC 4962  
Qy 541 ThrHisIleAspAlaHisPheLeuSerGlnThrLysGlnSerGlyGluAsnLeuProTyr 560  
Db 4963 ACTCATATAGATGCCACTTTCATCCAGACAAGCAGAGTGGGGAGAACCTTCTCTTAC 5022  
Qy 561 LeuValAlaTyrGlnAlaThrValCysAlaArgAlaGlnAlaProProSerTrpAsp 580  
Db 5023 CTGGTAGGTACCAAGCCACCGTGTGCGCTAGGCTCAAGCCCTCCCCCATCGTGGGAC 5082  
Qy 581 GlnMetTrpLysCysLeuIleArgLeuLysProThrLeuHisGlyProThrProLeuLeu 600  
Db 5083 CAGATGTGGAAGTGTGATTCGCTCAAGCCCACTCCATGGGCCAACACCCCTGCTA 5142  
Qy 601 TyrArgLeuGlyAlaValGlnAsnGluValThrLeuThrHisProValThrLysTyrIle 620  
Db 5143 TACAGACTGGGCGCTGTTCAGATGAAATCACCTGACGACCCAGTCACCAATAACATC 5202  
Qy 621 MetThrCysMetSerAlaAspLeuGluValValThrSerThrTrpValLeuValGlyGly 640  
Db 5203 ATGACATGCACTGTCGGCGGACCTGGAGGTGCTCAGAGACCTGGGTGCTGTTGGCGGC 5262  
Qy 641 ValLeuAlaLeuAlaAlaTyrCysLeuSerThrGlyCysValValIleValGlyArg 660  
Db 5263 GTCTGGCTGCTTTGGCGCGTATTGCTGTCAACAGGCTGCGTGTGTCATAGTGGCAGG 5322  
Qy 661 IleValLeuSerGlyLysProAlaIleIleProAspArgGluValLeuTyrArgGluPhe 680  
Db 5323 GTGCTCTGTTCGGGGAAGCGGCAATCATCTGACAGGGAAGTCTCTACCGAGAGTTC 5382  
Qy 681 AspGluMetGluGluCys 686  
Db 5383 GATGAGATGGAAGAGTGC 5400  
RESULT 11  
AAQ10566  
ID AAQ10566 standard; DNA; 9185 BP.  
XX AAQ10566;  
AC AAQ10566;  
XX 25-MAR-2003 (revised)  
DT 29-APR-1991 (first entry)  
XX Hepatitis C virus strain 1 DNA.  
DE Hepatitis C virus; HCV-1; non-A, non-B hepatitis; HCV antigen;  
XX viral infections; ss.  
KW Hepatitis C virus.  
OS OS  
XX EP414475-A.  
PN 27-FEB-1991.  
XX 21-AUG-1990; 90EP-00309120.  
PF 25-AUG-1989; 89US-00398667.  
XX (CHIR ) CHIRON CORP.  
PA



XX	Weiner AJ, Steimer KS;
FI	WPI; 1991-059670/09.
XX	Cell lines infected with hepatitis C virus - are used as source of
DR	antigens for detection of HCV antibodies, for vaccines, and for screening
XX	anti-viral agents.
PT	Disclosure; Fig 1; 24pp; English.
PT	This is a hepatitis C virus (HCV) composite cDNA sequence, deduced using
XX	overlapping clones. A compsn. confg. the antigenic protein encoded by
CC	this sequence is useful for detecting anti-HCV anti- bodies (Abs) and for
CC	screening an agent which inhibits HCV replic- ation. A cell line infected
CC	with this virus can be used as a source of antigens. The antigen is
CC	useful for preparing vaccines for treating viral infections. See also
CC	AQ10567. (Updated on 25-MAR-2003 to correct PA field.)
XX	SQ Sequence 9185 BP; 1849 A; 2790 C; 2608 G; 1938 T; 0 U; 0 Other;
Alignment Scores:	
Pred. No.:	2,48e-241 Length: 9185
Score:	3574.00 Matches: 672
Percent Similarity:	99.42% Conservative: 10
Best Local Similarity:	97.96% Mismatches: 4
Query Match:	98.78% Indels: 0
DB:	Gaps: 0
US-09-930-591-2 (1-686) x AAQ10566 (1-9185)	
Qy	1 MetAlaProIleThrAlaTyrAlaGlnGlnThrArgGlyLeuLeuGlyCysAileIleThr 20
Db	::::
3395	CTGGCGCCCATCACGCCGTACGCCAGCAGACAAGGGGCTCTTAGGGTGATAATCAACC 3454
Qy	21 SerLeuThrClYArgAspIysAnGlnValGluGlyGluValGlnIleValSerThAla 40
Db	AGCCTAACTTGGCGGGGACAAAACCAAGTGGAGGGTGAGGTCCAGATTGTGTCAACTGCT 3514
Qy	41 AlaGlnThrPheLeuAlaThrCysIleAnGlyValCysTrpThrValTyHisGlyVala 60
Db	GCCAAACCTTCCTGGCAACGTGCATCAATGGGGTGTCTGACATGCTACACGGGGCCC 3574
Qy	61 GlyThrArgThrIleAlaSerProIySglyProValIleGlnMetTyrThrAsnValasp 80
Db	GGAACGAGGACCATCGCTCACCCAGGGTCTGTGCATCCAGATGTATACCAATGTAGAC 3634
Qy	81 GlNAspleuValGlyTrpProAlaProGlnGlyAlaArGserLeuThrProCyThrCys 100
Db	CAGAAGCTTTGTGGCTTGGCGCGCTCGCAAGGTAGCCGCTCATTGACACCCCTGCACCTGC 3694
Qy	101 GlySerSerAspleuTyrLeuValThrArgHisAlaAspValIleProValArgArg 120
Db	GGCTCTCGGACCTTTACTCTGTGCAGAGCAGCCGAGTGCATTTCCCGTGGCGGGCGG 3754
Qy	121 GlyASpglyArgGlySerLeuSerProArgProIleSerTyrLeuLysGlySerSer 140
Db	GGTGATAGCAGGGCAGCTGTGTGCCCCGCCCATTTCTTACTTGAAGGGTCTCTCG 3814
Qy	141 GlyGlyProLeuLeuCysProAlaGlyHisAlaValAlGlyIlePheArgAlaAlaValCys 160
Db	GGGGGTCCGTGTGTGCCCCGGGGCAGCGCGTGGGCATATTTAGGGCCGCGGTGTGC 3874
Qy	161 ThrArgGlyValAlalysAlaValAspPheIleProValGluSerLeuGluThrThrMet 180
Db	ACCCGTGGAGTGGCTTAAGCGGTGGACTTTATTCCTGTGGAGAACCTTAGAGACAACCATG 3934
Qy	181 ArgSerProValPheSerAspAsnSerSerProProAlaValProGlnSerTyrGlnVal 200
Db	AGGTCCCGGTGTTCAGGATTAACCTCTCCACCAGTAGTGGCCCGCAGAGCTTCCAGGTG 3994
Qy	201 AlaHisLeuHisAlaProThrGlySerGlyIysSerThrIysValProAlaAlaTyrAla 220
3395	GCTCACTCTCATGCTCCACAGCAGCGGCAAAACCAAGGTCCCGGCTGCATATATGCA 4054
221	AlaGlnGlyTyrLysValLeuValLeuAsnProSerValAlaAlaThrMetGlyPheGly 240
4055	GCTCAGGGCTATAAGGTGCTAGTACTCAACCCCTCTGTGCTGCAACACTGGGCTTTGGT 4114
241	AlaTyrMetSerLysAlaHisGlyIleAspProAsnIleArgThrGlyValArgThrIle 260
4115	GCTTACATGTCNAAGGCTCATGGATCGATCCTAAACATCAGGACCGGGTGAGAACAAATT 4174
261	ThrThrGlySerProIleThrTyrSerThrTyrGlyLysPheIleuAlaAspGlyGlyCys 280
4175	ACCATGGCAGCCCCCATCAGTACTCCACTACGGCAAGTTCTTTCGCGACGGGGGTGC 4234
281	SerGlyGlyValAtyAspIleIleIleCysAspGluCysHisSerThrAspAlaThrSer 300
4235	TGGGGGGGCGCTTATGACATAATAATTGTGACGAGTGCCACTCCACGGATGCCACATCC 4294
301	IleLeuGlyIleGlyThrValLeuAspGlnAlaGluThrAlaGlyAlaArgLeuThrVal 320
4295	ATCTTGGGCATCGGCACCTGCTTGACCAAGCAGAGACTGCGGGGGGAGACTGGTTGTG 4354
321	LeuAlaThrAlaThrProProGlySerValThrValProHisProAsnIleGluGluVal 340
4355	CTCGCCACGGCACCCTCCCGGCTCCGTCACTGTGCCCATCCCAACATCGAGGAGTT 4414
341	AlaLeuSerThrThrGlyGluIleProPheTyrGlyLysAlaIleProLeuGluAlaIle 360
4415	GCTCTGTCCAACCGGAGAGATCCCTTTTTACGGCAAGGTATCCCCCTCGAAGTANATC 4474
361	LysGlyGlyArgHisLeuIlePheCysHisSerIysLysLysCysAspGluLeuAlaAla 380
4475	AAGGGGGGAGACATCTCATCTTCTGTCTTCAAAGAAGAAGTSCGACGAACTCGCGCA 4534
381	LysLeuValAlaLeuGlyValAsnaValAlaTyrTyrArgGlyLeuAspValSerVal 400
4535	AAGCTGGTCGATTTGGGCATCAATGCCGTGCTTACCCGGCTCTTGACGTGTCGCTC 4594
401	IleProThrSerGlyVasplValValAlaThrAspAlaLeuMetThrGlyPheThr 420
4595	ATCCCGACGAGGGGAGTGTGTGTGTGGCAACCGATGCCCTCATGACCGGCTATACC 4654
421	GlyAspPheAspSerValIleAspCysAsnThrCysValThrGlnThrValAspPheSer 440
4655	GGCGACTTCGATCGGTGATAGCTGCAATACGTGTGTCAACCCAGACAGTCGATTTTCA 4714
441	LeuAspProThrPheThrIleGluThrIleThrLeuProGlnAspAlaValSerArgThr 460
4715	CTTGACCTTACCTTCCACATTCAGCAATCAACGCTCCCGCAGGATGTCTGTCCCGCACT 4774
461	GlnArgArgGlyArgThrGlyArgGlyLysProGlyIleTyrArgPheValAlaProGly 480
4775	CAACCTCGGGGAGACTTGGCAGGGGGAAGCAGGCATCTACAGATTTTGGCACCGGGG 4834
481	GluArgProSerGlyMetPheAspSerSerValLeuCysGluCysTyrAspAlaGlyCys 500
4835	GAGCGCCCTCCGCGATGTTTCGACTCGTCCGTCTCTGTGAGTGTATGACGAGGCTGT 4894
501	AlaTyrTyrGluLeuThrProAlaGluThrThrValArgLeuArgAlaTyrMetAsnThr 520
4895	GCTTGGTATGAGCTCACGCCCGCAGACTACAGTTTAGGCTACGAGCGTACATGAACACC 4954
521	ProGlyLeuProValCysGlnAspHisLeuGluPheTrpGluGlyValPheThrGlyLeu 540
4955	CCGGGGCTTCCCGTGGCAGGACCATCTTGAAATTTTGGAGGGGCGTCTTACAGGCCTC 5014
541	ThrHisIleAspAlaHisPheLeuSerGlnThrLysGlnSerGlyGluAsnLeuProTyr 560
5015	ACTCATATAGATGCCCATCTTCTATCCAGACAAAGCAGAGTGGGAGAACCTTCTCTTAC 5074
561	LeuValAlaTyrGlnAlaThrValCysAlaArgAlaGlnAlaProProSerTrpAsp 580
5075	CTGGTAGCGTACCAAGCCACCGTGTGGCTAGGGCTCAAGCCCTCCCATCTGCTGGGAC 5134

QY 501 GlnMetTrpLysCysLeuIleAArgLeuLysProThrLeuHisGlyProThrProLeuLeu 600  
DB 5135 CAGATGTGGAAGTGTGTTGATTGCGCTCAAGCCACCTCCATGGGCCAACACCCCTGCTA 5194  
QY 601 TyrArgLeuGlyAlaValGlnAsnGluValThrLeuThrHisProValThrLysTrile 620  
DB 5195 TACAGACTGGCGCTGTTGAGATGAAATACACCTGACGACCCAGTCCACAAATACATC 5254  
QY 621 MetThrCysMetSerAlaAspLeuGluValValThrSerThrTrpValLeuValGlyGly 640  
DB 5255 ATGACATGCATGTGGCGGACCTGGAGGTGTCACGAGCACCTGGGTGCTGTTGGCGGC 5314  
QY 641 ValLeuAlaAlaLeuAlaTyrCysLeuSerThrGlyCysValValLleValGlyArg 660  
DB 5315 GTCTCTGGCTGTTGGCGCGCTATTGCTGTCAACAGGCTGCTGTGTCATAGTGGCAGG 5374  
QY 661 IleValLeuSerGlyLysProAlaIleProArgGluValLeuValThrArgGluPhe 680  
DB 5375 GTCTGTCTGTCCGGAAGCCGCAATCATCTGACAGGAAGTCTCTACCGAGAGTTC 5434  
QY 681 AspGluMetGluGluCys 686  
DB 5435 GATGAGATGGAGAGTGC 5452

## RESULT 12

AAA75297  
ID AAA75297 standard; cDNA; 9185 BP.

XX  
AC AAA75297;

XX  
DT 15-JAN-2001 (first entry)

XX  
DE Sense strand of HCV encoding a polyprotein.

XX  
KW Hepatitis C virus; HCV; antisense polynucleotide; polyprotein;

XX  
KW viral infectivity; viral replication; ds.

XX  
OS Hepatitis C virus.

XX  
FH Key Location/Qualifiers

FT CDS 320..9184

FT /\*tag= a

FT /note= "partial sequence; no termination codon given"

XX  
EP1034785-A2.

XX  
PD 13-SEP-2000.

XX  
PF 16-MAR-1990; 2000EP-00109602.

XX  
PR 17-MAR-1989; 89US-00325338.

XX  
PR 20-APR-1989; 89US-00341334.

XX  
PR 18-MAY-1989; 89US-00355002.

XX  
PR 16-MAR-1990; 90EP-00302866.

XX  
PA (CHIR ) CHIRON CORP.

XX  
PI Houghton M, Choo Q, Kuo G;

XX  
DR WPI; 2000-566891/53.

XX  
DR P-PSDB; AAB18541.

XX  
PT Novel composition comprising a hepatitis C virus antisense polynucleotide

XX  
PT which is complementary to or corresponds to a sense strand of the virus

XX  
PT genome, and selectively hybridizes to it.

XX  
PS Example; Fig 17; 75pp; English.

XX  
CC The specification describes a pharmaceutical composition which comprises

XX  
CC a hepatitis C virus (HCV) antisense polynucleotide. The HCV is

XX  
CC characterized by a positive stranded RNA genome which has 40% homology at

XX  
CC the polypeptide level to a HCV polyprotein. The antisense polynucleotide

CC binds to cellular polynucleotides which enhance and/or are required for  
CC viral infectivity, replicative ability or chronicity. The antisense  
CC polynucleotides may also be designed to bind with high specificity, to be  
CC of increased stability, to be stable and to have low toxicity. The  
CC composition also comprises an agent which causes viral RNA to be  
CC inactive. The composition is used for preventing HCV replication in a  
CC system. The present sequence represents a novel HCV cDNA sequence, which  
CC is used in the course of the invention

XX  
SQ Sequence 9185 BP; 1849 A; 2790 C; 2608 G; 1938 T; 0 U; 0 Other;

## Alignment Scores:

Pred. No.: 2,48e-241 Length: 9185  
Score: 3574.00 Matches: 672  
Percent Similarity: 99.42% Conservative: 10  
Best Local Similarity: 97.96% Mismatches: 4  
Query Match: 98.78% Indels: 0  
DB: 3 Gaps: 0

US-09-930-591-2 (1-686) x AAA75297 (1-9185)

QY 1 MetAlaProIleThrAlaTyrAlaGlnGlnThrArgGlyLeuLeuGlyCysIleIleThr 20

DB 3395 CTGGCGGCCCATCAGCGGTACGCCCCAGACAGCAAGGGCGCTCTTAGGGTGATATACACC 3454

QY 21 SerLeuThrGlyArgAspLysAsnGlnValGluGlyGluValGlnIleValSerThrAla 40

DB 3455 AGCCTAACTGGCGGGACAAAACCAAGTGAGGGTGCAGATTGTGTCAACTGCT 3514

QY 41 AlaGlnThrPheLeuAlaThrCysIleAsnGlyValCysTrpThrValTyrHisGlyAla 60

DB 3515 GCCCAAACTTCTCGGCAACGTGCATCAATGGGGTGTCTGAGCTGTCTACACGGGGGCC 3574

QY 61 GlyThrArgThrIleAlaSerProLysGlyProValIleGlnMetTyrThrAsnValAsp 80

DB 3575 GGAACGAGGACCATCGCGTCACCCAGGGTCTGTCTATCCAGATGTATACCAATGTAGAC 3634

QY 81 GlnAspLeuValGlyTrpProAlaProGlnGlnAlaArgSerLeuThrProCysThrCys 100

DB 3635 CAAGACCTTGTGGGCTGGCGGCTCCGCAAGTAGCCGCTCATTTGACACCCCTGCACATTGC 3694

QY 101 GlySerSerAspLeuThrLeuValThrArgHisAlaAspValIleProValArgArg 120

DB 3695 GGCTCTCGGACCTTTACCTGGTTCAGAGGCACCCGATGTCAATCCGTCGCGCGCGCG 3754

QY 121 GlyAspGlyArgGlySerLeuLeuSerProArgProIleSerTyrLeuLysGlySerSer 140

DB 3755 GGTGATAGCAGGGGACGCTGTCTGCGCCCGGCCCATTTCTCTACTTGAAGGCTCTCTCG 3814

QY 141 GlyGlyProLeuLeuCysProAlaGlyHisAlaValGlyIlePheArgAlaAlaValCys 160

DB 3815 GGGGTCGCTGTTGTGCCCGCGGGGACGCGTGGGCATATTTAGGGCGCGGTGTGC 3874

QY 161 ThrArgGlyValAlaLysAlaValAspPheIleProValGluSerLeuGluThrThrMet 180

DB 3875 ACCGTGGAGTGGCTAAGGGGTGGACTTTATCTCTGTGAGAGACCTAGACACCAACATG 3934

QY 181 ArgSerProValPheSerAspAsnSerSerProAlaValProGlnSerTyrGlnVal 200

DB 3935 AGGTCCCGGTGTTACGGATACTCTCTCCACAGTAGTGCCCGAGCTTCCAGGTG 3994

QY 201 AlaHisLeuHisAlaProThrGlySerGlyLysSerThrLysValProAlaAlaTyrAla 220

DB 3995 GCTCACCCTCCATGCTCCACAGGCGGCAAGACCAAGGTCCTGGCTGCATATGCA 4054

QY 221 AlaGlnGlyTyrLysValLeuValLeuAsnProSerValAlaAlaThrMetGlyPheGly 240

DB 4055 GCTCAGGGCTATAAGGTGCTAGTACTCAACCCCTCTGTGTGTGCAACACTGGGCTTTGGT 4114

QY 241 AlaTyrMetSerLysAlaHisGlyIleAspProAsnIleArgThrGlyValArgThrIle 260

DB 4115 GCTTACATGTCCAGGCTCATGGATCGATCCTTAACATCAGGACCGGGGTGAGAACAAAT 4174



Qy	1	MetAlaProIleThrAlaTyrAlaGlnGlnThrArgGlyLeuLeuGlyCysIleIleThr	20
Db	3395	CTGGCGCCCATACAGGGCGTAGCCCCAGCAGACAAGGGGCGCTCTCTAGGGTGCATATACACC	3454
Qy	21	SerLeuThrGlyArgAspLysAsnGlnValGluGlyGluValGlnIleValSerThrAla	40
Db	3455	AGCCTNACTGGCCGGGACAAAAACCAAGTGAGGGTGAGGTCCAGATTGTGTCAACTGCT	3514
Qy	41	AlaGlnThrPheLeuAlaThrCysIleAsnGlyValCysTrpThrValTyrHisGlyAla	60
Db	3515	GCCCAAACTTCTTGGCAACGTGCATCAATGGGTGTGCTGGACTGTCTACCAACGGGCGC	3574
Qy	61	GlyThrArgThrIleAlaSerProIysGlyProValIleGlnMetTyrThrAsnValAsp	80
Db	3575	GGAAACGAGACCATCGCGTCAACCAAGGGTCTCTGTCATCCAGATGTATACCAATGTAGAC	3634
Qy	81	GlnAspLeuValGlyTrpProAlaProGlnGlyAlaArgSerLeuThrProCysThrCys	100
Db	3635	CAAGACCTTGTGGGTGGCCCGCTCCGCAAGTAGCGCTCATTTGACACCTCGACCTTGC	3694
Qy	101	GlySerSerAspLeuTyrLeuValThrArgHisAlaAspValIleProValArgArg	120
Db	3695	GGCTCTCGGACCTTACTTGGTCAAGGACACCGCATGTCAATCCGTCGCGCGCGG	3754
Qy	121	GlyAspGlyArgGlySerLeuLeuSerProArgProIleSerTyrLeuLysGlySerSer	140
Db	3755	GGTGATAGCAGGGCGAGCTCTGTGCGCCCGGCCCATTTCTACTTTGAAAGGCTCTCTCG	3814
Qy	141	GlyGlyProLeuLeuCysProAlaGlyHisAlaValGlnIlePheArgAlaAlaValCys	160
Db	3815	GGGGGTCCGCTGTGTGCCCCCGCGGGCACGCCGTGGGCAATTTAGCGCGCGGTGTGC	3874
Qy	161	ThrArgGlyValAlaLysAlaValAspPheIleProValGluSerLeuGluThrThrMet	180
Db	3875	ACCCGTGGAGTGGCTAAGGCGGTGGACTTTATCTCTGTGGAGAACCTTAGAGACAACCATG	3934
Qy	181	ArgSerProValPheSerAspAsnSerSerProProAlaValProGlnSerTyrGlnVal	200
Db	3935	AGGTCCCGGGTGTTCAGGGATAACTCTCTCCACAGTAGTGCCTCCAGAGCTTCCAGGTG	3994
Qy	201	AlaHisLeuHisAlaProThrGlySerGlyLysSerThrLysValProAlaAlaTyrAla	220
Db	3995	GCTCACCTCCATGTCTCCACAGGACGGGCAAAAGCACCAAGGTCCCGGCTGCATATGCA	4054
Qy	221	AlaGlnGlyTyrLysValLeuValLeuAsnProSerValAlaAlaThrMetGlyPheGly	240
Db	4055	GCTCAGGGCTATAGGTGCTAGTACTCAACCCCTCTGTTGCTGCTCAACACCTGGGGCTTGGT	4114
Qy	241	AlaTyrMetSerLysAlaHisGlyIleAspProAsnIleArgThrGlyValArgThrIle	260
Db	4115	GCTTACATGTCCAAGGCTCATGGGATCGATCCTAAATCAGGACCGGGGTGAGAACAAATT	4174
Qy	261	ThrThrGlySerProIleThrTyrSerThrTyrGlyLysPheLeuAlaAspGlyGlyCys	280
Db	4175	ACCACTGGACGCCCATCAGTACTCACCTACGGCAAGTTCTCTTGCACGCGCGGGTGC	4234
Qy	281	SerGlyGlyAlaTyrAspIleIleCysAspGluCysHisSerThrAspAlaThrSer	300
Db	4235	TCGGGGGGCGCTTATGACATATATTTGTGACAGGTGCCACTCCACGGATGCCATATCC	4294
Qy	301	IleLeuGlyIleGlyThrValLeuAspGlnAlaGluThrAlaGlyAlaArgLeuThrVal	320
Db	4295	ATCTTGGGCATCGGCACTGTCTTTGACCAAGCAGAGACTGCGGGGGCGAGACTGGTTGTG	4354
Qy	321	LeuAlaThrAlaThrProProGlySerValThrValProHisProAsnIleGluGluVal	340
Db	4355	CTCGCCACCGGACCCCTTCGCGGCTCGGTCACTGTGCGCCCATCTCCAACTCAGGAGGTT	4414
Qy	341	AlaLeuSerThrThrGlyGluIleProPheTyrGlyLysAlaIleProLeuGluAlaIle	360
Db	4415	GCTCTGTCCACCCGAGAGATCCCTTTTACGGCAAGGCTATCTCCCTTCGAAGTAACT	4474
Qy	361	LysGlyGlyArgHisLeuIlePheCysHisSerLysLysLysCysAspGluLeuAlaAla	380

## RESULT 14

AAT12710

ID AAT12710 standard: cDNA: 9401 BP.  
 AAT12710

XX

AC AAT12710;

XX

Db	4475	AAGGGGGGAGACATCTCATCTTCTGTCATTTCAAGAGAAGTGGCGACGAACCTCGCGCA	4534
Qy	381	lysLeuValAlaLeuGlyValAsnAlaValAlaTyrTyrArgGlyLeuAspValSerVal	400
Db	4535	AAGCTGGTCGATTTGGGCATCAATGCGTGGCCTACTACCGCGGTCTTGACGTGTCGTC	4594
Qy	401	IleProThrSerGlyAspValValValAlaThrAspAlaLeuMetThrGlyPheThr	420
Db	4595	ATCCGACCAAGCGCGATGTTGTGCTGTGGCAACCGATGCCCTCATATGACCGGCTATACC	4654
Qy	421	GlyAspPheAspSerValIleAspCysAsnThrCysValThrGlnThrValAspPheSer	440
Db	4655	GGCGACTTCGACTCGGTGATAGACTGCAATACGTGTGTACCCAGACAGTCGATTTTCAGC	4714
Qy	441	LeuAspProThrPheThrIleGluThrIleThrLeuProGlnAspAlaValSerArgThr	460
Db	4715	CTTGACCTACCTTACCATTTGAGACAATCACGCTCCCCAGAGATGTGTCTCCCGCACT	4774
Qy	461	GlnArgArgGlyArgThrGlyArgGlyLysProGlyIleIleTyrArgPheValAlaProGly	480
Db	4775	CAACGTCGGGGCAGGACTGCGACGGGAGCCAGGCAATCTACAGATTTGTGGCACCGGGG	4834
Qy	481	GluArgProSerGlyMetPheAspSerSerValLeuCysGluCysTyrAspAlaGlyCys	500
Db	4835	GAGCGCCCTCCGCGATGTTCGACTCGTCGTCCTCTGTGAGTGTATGACGCAGGCTGT	4894
Qy	501	AlaTrpTyrGluLeuThrProAlaGluThrThrValArgLeuArgAlaTyrMetAsnThr	520
Db	4895	GCTTGGTATGAGCTACGCCCGCGAGACTACAGTTAGGCTACGAGCGTACATGAACACC	4954
Qy	521	ProGlyLeuProValCysGlnAspHisLeuGluPheTrpGluGlyValPheThrGlyLeu	540
Db	4955	CCGGGGCTTCCCGTGTGCCAGGACCATCTTGAATTTTGGAGGGCGCTCTTTACAGGCCTC	5014
Qy	541	ThrHisIleAspAlaHisPheLeuSerGlnThrLysGlnSerGlyGluAsnLeuProTyr	560
Db	5015	ACTCATATAGATCCCATTTTCTATCCAGACAAAGCAGATGGGAGAAACCTTCCTCTAC	5074
Qy	561	LeuValAlaTyrGlnAlaThrValCysAlaArgAlaGlnAlaProProSerTrpAsp	580
Db	5075	CTGGTAGGTACCAAGCCACCGTGTGGCTTAGGGCTCAAGCCCTCCCCCATCGTGGAC	5134
Qy	581	GlnMetTrpLysCysLeuIleArgLeuLysProThrLeuHisGlyProThrProLeuLeu	600
Db	5135	CAGATGTGGAAGTGTGTTGATTTCGCTCAAGCCACCGCTCCATGGGCAACACCCCTGCTA	5194
Qy	601	TyrArgLeuGlyAlaValGlnAsnGluValThrLeuThrHisProValThrLysTyrIle	620
Db	5195	TACAGACTGGGCGCTGTTTCAGATGAATACCTCGACGCCCCAGTGCACCAANTACATC	5254
Qy	621	MetThrCysMetSerAlaAspLeuGluValValThrSerThrTrpValLeuValGlyGly	640
Db	5255	ATGACATGCTATGCGGCCGACCTGGAGTGTGTCAACGAGCACCTGGTGTCTGTTGGCGC	5314
Qy	641	ValLeuAlaIleAlaLeuAlaTyrCysLeuSerThrGlyCysValIleValGlyArg	660
Db	5315	GTCTCGTCTGTTGGCGCGTATTCCTGTCAACAGGGCTGGTGTGTATGATGGGCAGG	5374
Qy	661	IleValLeuSerGlyLysProAlaIleIleProAspArgGluValLeuTyrArgGluPhe	680
Db	5375	GTGCTCTGTTCGGGAAGCCGGCAATCATACCTGACAGGGAGTCTCTTACCGAGAGTTC	5434
Qy	681	AspGluMetGluGluCys	686
Db	5435	GATGAGATGGAAGTGC	5452

## RESULT 14

REC'D 17  
AAT12710

ID AAT12710 standard: cDNA: 9401 BP.  
 AAT12710

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65  
66  
67  
68  
69  
70  
71  
72  
73  
74  
75  
76  
77  
78  
79  
80  
81  
82  
83  
84  
85  
86  
87  
88  
89  
90  
91  
92  
93  
94  
95  
96  
97  
98  
99  
100  
101  
102  
103  
104  
105  
106  
107  
108  
109  
110  
111  
112  
113  
114  
115  
116  
117  
118  
119  
120  
121  
122  
123  
124  
125  
126  
127  
128  
129  
130  
131  
132  
133  
134  
135  
136  
137  
138  
139  
140  
141  
142  
143  
144  
145  
146  
147  
148  
149  
150  
151  
152  
153  
154  
155  
156  
157  
158  
159  
160  
161  
162  
163  
164  
165  
166  
167  
168  
169  
170  
171  
172  
173  
174  
175  
176  
177  
178  
179  
180  
181  
182  
183  
184  
185  
186  
187  
188  
189  
190  
191  
192  
193  
194  
195  
196  
197  
198  
199  
200  
201  
202  
203  
204  
205  
206  
207  
208  
209  
210  
211  
212  
213  
214  
215  
216  
217  
218  
219  
220  
221  
222  
223  
224  
225  
226  
227  
228  
229  
230  
231  
232  
233  
234  
235  
236  
237  
238  
239  
240  
241  
242  
243  
244  
245  
246  
247  
248  
249  
250  
251  
252  
253  
254  
255  
256  
257  
258  
259  
260  
261  
262  
263  
264  
265  
266  
267  
268  
269  
270  
271  
272  
273  
274  
275  
276  
277  
278  
279  
280  
281  
282  
283  
284  
285  
286  
287  
288  
289  
290  
291  
292  
293  
294  
295  
296  
297  
298  
299  
300  
301  
302  
303  
304  
305  
306  
307  
308  
309  
310  
311  
312  
313  
314  
315  
316  
317  
318  
319  
320  
321  
322  
323  
324  
325  
326  
327  
328  
329  
330  
331  
332  
333  
334  
335  
336  
337  
338  
339  
340  
341  
342  
343  
344  
345  
346  
347  
348  
349  
350  
351  
352  
353  
354  
355  
356  
357  
358  
359  
360  
361  
362  
363  
364  
365  
366  
367  
368  
369  
370  
371  
372  
373  
374  
375  
376  
377  
378  
379  
380  
381  
382  
383  
384  
385  
386  
387  
388  
389  
390  
391  
392  
393  
394  
395  
396  
397  
398  
399  
400  
401  
402  
403  
404  
405  
406  
407  
408  
409  
410  
411  
412  
413  
414  
415  
416  
417  
418  
419  
420  
421  
422  
423  
424  
425  
426  
427  
428  
429  
430  
431  
432  
433  
434  
435  
436  
437  
438  
439  
440  
441  
442  
443  
444  
445  
446  
447  
448  
449  
450  
451  
452  
453  
454  
455  
456  
457  
458  
459  
460  
461  
462  
463  
464  
465  
466  
467  
468  
469  
470  
471  
472  
473  
474  
475  
476  
477  
478  
479  
480  
481  
482  
483  
484  
485  
486  
487  
488  
489  
490  
491  
492  
493  
494  
495  
496  
497  
498  
499  
500  
501  
502  
503  
504  
505  
506  
507  
508  
509  
510  
511  
512  
513  
514  
515  
516  
517  
518  
519  
520  
521  
522  
523  
524  
525  
526  
527  
528  
529  
530  
531  
532  
533  
534  
535  
536  
537  
538  
539  
540  
541  
542  
543  
544  
545  
546  
547  
548  
549  
550  
551  
552  
553  
554  
555  
556  
557  
558  
559  
560  
561  
562  
563  
564  
565  
566  
567  
568  
569  
570  
571  
572  
573  
574  
575  
576  
577  
578  
579  
580  
581  
582  
583  
584  
585  
586  
587  
588  
589  
590  
591  
592  
593  
594  
595  
596  
597  
598  
599  
600  
601  
602  
603  
604  
605  
606  
607  
608  
609  
610  
611  
612  
613  
614  
615  
616  
617  
618  
619  
620  
621  
622  
623  
624  
625  
626  
627  
628  
629  
630  
631  
632  
633  
634  
635  
636  
637  
638  
639  
640  
641  
642  
643  
644  
645  
646  
647  
648  
649  
650  
651  
652  
653  
654  
655  
656  
657  
658  
659  
660  
661  
662  
663  
664  
665  
666  
667  
668  
669  
670  
671  
672  
673  
674  
675  
676  
677  
678  
679  
680  
681  
682  
683  
684  
685  
686  
687  
688  
689  
690  
691  
692  
693  
694  
695  
696  
697  
698  
699  
700  
701  
702  
703  
704  
705  
706  
707  
708  
709  
710  
711  
712  
713  
714  
715  
716  
717  
718  
719  
720  
721  
722  
723  
724  
725  
726  
727  
728  
729  
730  
731  
732  
733  
734  
735  
736  
737  
738  
739  
740  
741  
742  
743  
744  
745  
746  
747  
748  
749  
750  
751  
752  
753  
754  
755  
756  
757  
758  
759  
760  
761  
762  
763  
764  
765  
766  
767  
768  
769  
770  
771  
772  
773  
774  
775  
776  
777  
778  
779  
780  
781  
782  
783  
784  
785  
786  
787  
788  
789  
790  
791  
792  
793  
794  
795  
796  
797  
798  
799  
800  
801  
802  
803  
804  
805  
806  
807  
808  
809  
810  
811  
812  
813  
814  
815  
816  
817  
818  
819  
820  
821  
822  
823  
824  
825  
826  
827  
828  
829  
830  
831  
832  
833  
834  
835  
836  
837  
838  
839  
840  
84

AC AAT12710;

2000



```
Db 4737 CTTGACCCCTACCTTCCACCATTCAGACAAATCAGGCTCCCGGAGTGTCTCCCGGCACT 4796
Qy 461 GlnArgArgGlyArgThrGlyArgGlyLysProGlyIleTyrArgPheValAlaProGly 480
Db 4797 CAACGTGGGGCAGGACTGGCAGGGGGAAGCCAGGCATCTACAGATTTGTGGCACC GG 4856
Qy 481 GluArgProSerGlyMetPheAspSerSerValLeuCysGluCysTyrAspAlaGlyCys 500
Db 4857 GAGCGCCCTCCGGCATGTTGACTCGCTCGCTCTGTGAGTGTATGACGCGAGGCTGT 4916
Qy 501 AlaTyrTyrGluLeuThrProAlaGluThrThrValArgLeuArgAlaTyrMetAsnThr 520
Db 4917 GCTTGATAGTCTCACCCGCCAGACTACAGTTAGCTACGAGCGTACATGAAACACC 4976
Qy 521 ProGlyLeuProValCysGlnAspHisLeuGluPheTyrGluGlyValPheThrGlyLeu 540
Db 4977 CGGGGCTTCGCTGTGTCAGGACCATCTTGAATTTTGGAGGGGCTCTTTACAGGCTC 5036
Qy 541 ThrHisIleAspAlaHisPheLeuSerGlnThrLysGlnSerGlyGluAsnLeuProTyr 560
Db 5037 ACTCATATAGATGCCCACTTTCTATCCAGACAAAGCAGAGTGGGGGAGAACCTTCTTAC 5096
Qy 561 LeuValAlaTyrGlnAlaThrValCysAlaArgAlaGlnAlaProProSerTyrAsp 580
Db 5097 CTGGTAGCGTACCAGCCACCGTGTGCGCTAGGCGCTCAAGCCCTCCCGGATCGTGGGAC 5156
Qy 581 GlnMetTrpLysCysLeuIleArgLeuLysProThrLeuHisGlyProThrProLeuLeu 600
Db 5157 CAGATGTGGAGTGTGTTGATTCGCTCAAGCCACCTCCATGGGCCAACACCTCTGCTA 5216
Qy 601 TyrArgLeuGlyAlaValGlnAsnGluValThrLeuThrHisProValThrLysTyrIle 620
Db 5217 TACAGACTGGCGCTGTTTCAGAAATCAACCTGAGCCACCCAGTCAGCCCAATACATC 5276
Qy 621 MetThrCysMetSerAlaAspLeuValValThrSerThrTrpValLeuValGlyGly 640
Db 5277 ATGACATGATGTCGGCGACTGGAGTGTGTCAGAGCACCTGGGTCTCGTTGGGGC 5336
Qy 641 ValLeuAlaAlaLeuAlaAlaTyrCysLeuSerThrGlyCysValValIleValGlyArg 660
Db 5337 GTCTGTGCTGTTGGCGCGTATGCTGTCTACAGAGCTGCGTGTGTCATAGTGGGAGG 5396
Qy 661 IleValLeuSerGlyLysProAlaIleIleProAspArgGluValLeuTyrArgGluPhe 680
Db 5397 GTCGTCTGTTCGGGAGCGGCAATCATCTGACAGGAAGTCTCTACCGAGAGTTC 5456
Qy 681 AspGluMetGluGluCys 686
Db 5457 GATGAGTGGAGAGTGC 5474
RESULT 15
AAT99981
ID AAT99981 standard; DNA; 9401 BP.
XX
AC AAT99981;
XX
DT 25-MAR-2003 (revised)
DT 16-MAR-1998 (first entry)
XX
DE HCV polyprotein coding sequence.
XX
KW PCR primer; amplify; HCV; hepatitis c virus; antigen combination; NS3;
KW C domain; S domain; NS5; HCV polyprotein; anti-HCV antibody; detection;
KW NS4; ds.
XX
OS Hepatitis C virus.
XX
FH .Key Location/Qualifiers
FT CDS 342..9377
FT /*tag= a
XX
FN US5683864-A.
XX
```

```
PD 04-NOV-1997.
XX
XX 07-JUL-1992; 92US-00910760.
XX
PR 18-NOV-1987; 87US-00122714.
PR 30-DEC-1987; 87US-00139886.
PR 26-FEB-1988; 88US-00161072.
PR 06-MAY-1988; 88US-00191263.
PR 26-OCT-1988; 88US-00263584.
PR 14-NOV-1988; 88US-00271450.
PR 17-MAR-1989; 89US-00325338.
PR 20-APR-1989; 89US-00341334.
PR 21-APR-1989; 89US-00353896.
PR 18-MAY-1989; 89US-00355002.
PR 04-APR-1990; 90US-00504352.
XX
PA (CHIR ) CHIRON CORP.
XX
PI Kuo G, Houghton M, Choo Q;
XX
DR WPI; 1997-548976/50.
DR P-PSDB; AAW34480.
XX
PT Combination of three hepatitis C virus antigens - used for detection of
specific antibodies to diagnose infection.
XX
PS Disclosure; Col 25-46; 57pp; English.
XX
CC This sequence represents the Hepatitis C virus polyprotein coding
sequence. Fragments of this sequence can be amplified and used in the
combination of HCV antigens of the invention. The HCV antigen combination
comprises an antigen (Ag1) comprising the C domain (i.e. amino acids (aa)
1-120 of the HCV polyprotein), or its immunologically reactive fragment
containing at least 8 aa. It also comprises two additional antigens from
two different polyprotein domains, including at least 8 aa from the NS3,
NS4, S or NS5 domains of the polyprotein, corresponding, respectively, to
aa 1050-1640; 1640-2000; 120-400 and 2000-3011 of the HCV polyprotein.
CC Alternatively, Ag1 contains at least 8 aa from the 1-122 or 9-177 aa
regions of the HCV polyprotein. These antigen combinations are used
diagnostically to detect anti-HCV antibodies, using any standard
immunoassay format. These antigen combinations have a broader range of
reactivity with antibodies than any antigen individually. (Updated on 25-
MAR-2003 to correct PR field.)
XX
SQ Sequence 9401 BP; 1883 A; 2860 C; 2673 G; 1985 T; 0 U; 0 Other;
```

```
Alignment Scores:
Pred. NO.: 2.55e-241 Length: 9401
Score: 3574.00 Matches: 672
Percent Similarity: 99.42% Conservative: 10
Best Local Similarity: 97.96% Mismatches: 4
Query Match: 98.78% Indels: 0
DB: 2 Gaps: 0
```

US-09-930-591-2 (1-686) x AAT99981 (1-9401)

```
Qy 1 MetAlaProIleThrAlaTyrAlaGlnGlnThrArgGlyLeuLeuGlyCysIleIleThr 20
Db 3417 CTGGCGCCCATCGCGGTACGCCAGACAAAGGGGCTCTAGGTGTCATATCACC 3476
Qy 21 SerLeuThrGlyArgAspLysAsnGlnValGluGlyGluValGlnIleValSerThrAla 40
Db 3477 AGCTTAATCTGGCGGGACAAAACCAAGTGGAGGTGAGGTCCAGATTGTGTCAACTGCT 3536
Qy 41 AlaGlnThrPheLeuAlaThrCysIleAsnGlyValCysTrpThrValTyrHisGlyAla 60
Db 3537 GCCCAAAACCTTCTTGGCAACGTGCATCAATGGGGTGTCTGAGTGTCTACACGGGGCC 3596
Qy 61 GlyThrArgThrIleAlaSerProLysGlyProValIleGlnMetTyrThrAsnValAsp 80
Db 3597 GGAACGAGGACCATCGCGTCCACCAAGGGTCTCTCATCCAGATGTATACCAATGTAGAC 3656
Qy 81 GlnAspLeuValGlyTrpProAlaProGlnGlyAlaArgSerLeuThrProCysThrCys 100
```

```
Db 3657 CAAGACCTTGTGGCGCTGCCCGCTCCGCAAGGTAGCCGCTCATTTACACCCCTGCACTTGC 3716
Qy 101 GlySerSerAspLeuThrValThrArgHisAlaAspValIleProValArgArgArg 120
Db 3717 GGTCTCTCGGACCTTTACCTGGGTACAGGACGCGCATGTCATTCCCGTGCGCCGCGG 3776
Qy 121 GlyAspGlyArgGlySerLeuLeuSerProArgProIleSerTyrLeuLysGlySerSer 140
Db 3777 GGTGATAGAGGGGACGCTGTGTGCCCCCGGCCCATTTCTTACTTTGAAGGCTCTCTCG 3836
Qy 141 GlyGlyProLeuLeuCysProAlaGlyHisAlaValGlyIlePheArgAlaAlaValCys 160
Db 3837 GGGGTCCGCTGTGTGCCCCGGGGCACCGCTGGGCATATTTAGGGCGCGCGGTGC 3896
Qy 161 ThrArgGlyValAlaLysAlaValAspPheIleProValGluSerLeuGluThrThrMet 180
Db 3897 ACCCGTGGAGTGGCTAAGCGGGTGGACTTTATCCCTGTGGAGAACCTTAGAGACAACCATG 3956
Qy 181 ArgSerProValPheSerAspAsnSerSerProProAlaValProGlnSerTyrGlnVal 200
Db 3957 AGTCCCCGGTGTTCAGGATAACTCTCTCCACAGTAGTGCCCCAGAGCTTCCAGGTG 4016
Qy 201 AlaHisLeuHisAlaProThrGlySerGlyLysSerThrLysValProAlaIaIaIaVal 220
Db 4017 GCTCACCTCCATGCTCCACAGGCAGCGGCAAAAGCACCAAGTCCCGGCTGCATATGCA 4076
Qy 221 AlaGlnGlyTyrLysValLeuValLeuAsnProSerValAlaIaIaIaIaIaIaIaIaIa 240
Db 4077 GCTCAGGGCTATAAGGTGCTAGTACTCAACCCCTCTGTGCTGCAACACCTGGGCTTTGGT 4136
Qy 241 AlaTyrMetSerLysAlaHisGlyIleAspProAsnIleArgThrGlyValArgThrIle 260
Db 4137 GCTTACATGTCGAAGGCTCATGGATGCATCTTAACATCAGGACCGGGGTGAGAACATT 4196
Qy 261 ThrThrGlySerProIleThrTyrSerThrTyrGlyLysPheLeuAlaAspGlyGlyCys 280
Db 4197 ACCACTGGCAGCCCCCATCATGCTACTCCACCTACGGCAAGTTCTTGCAGCAGCGGGTGC 4256
Qy 281 SerGlyGlyAlaTyrAspIleIleIleCysAspGluCysHisSerThrAspAlaIaIaIa 300
Db 4257 TCGGGGGGCGCTTATGACATAATAATTTGTGACAGTGGCCACTCCACGGATGCCACATCC 4316
Qy 301 IleLeuGlyIleGlyThrValLeuAspGlnAlaGluThrAlaGlyAlaArgLeuThrVal 320
Db 4317 ATCTTGGGCATCGGCACTGCTTGTACCAAGCAGAGACTCGGGGGGCGAGACTGGTTGTG 4376
Qy 321 LeuAlaThrAlaThrProGlySerValThrValProHisProAsnIleGluGluVal 340
Db 4377 CTGCCACCGCCACCCCTCCGGGCTCCGTCCTCACTGTGCCCATCCCAACATCGAGGAGTT 4436
Qy 341 AlaLeuSerThrThrGlyGluIleProPheTyrGlyLysAlaIleProLeuGluAlaIle 360
Db 4437 GCTCTGTCCACCGGAGAGATCCCTTTTACGGCAAGGCTATCCCTCCGAAGTAATC 4496
Qy 361 LysGlyGlyArgHisLeuIlePheCysHisSerLysLysLysCysAspGluLeuAlaAla 380
Db 4497 AAGGGGGGAGACATCTCATCTCTGTCTATTCAAGNAGAGTGGCGAGCACTCGCGCA 4556
Qy 381 LysLeuValAlaLeuGlyValAsnAlaValAlaTyrTyrArgGlyLeuAspValSerVal 400
Db 4557 AAGCTGTGTCATTGGGCATCAATCGCGTCCGCTACTTACCGCGGTCTTTCAGCTGTCGTC 4616
Qy 401 IleProThrSerGlyAspValValValAlaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIa 420
Db 4617 ATCCCGACCGCGCGATGTGTGTCGTGCGCAACCGATGCCCTCATGACCGGTATACC 4676
Qy 421 GlyAspPheAspSerValIleAspCysAsnThrCysValThrGlnThrValAspPheSer 440
Db 4677 GGGGACTTCCACTCGGTCATAGACTGCATATGCTGTGTCTACCCAGACAGTCTGATTTACG 4736
Qy 441 LeuAspProThrPheThrIleGluThrIleThrLeuProGlnAspAlaValSerArgThr 460
```

```
Db 4737 CTTGACCCCTACCTTACCATTTGAGACAATCACGCTCCCCCAGGATGCTGTCTCCCGCACT 4796
Qy 461 GlnArgArgGlyArgThrGlyArgGlyLysProGlyIleTyrArgPheValAlaIaIaProGly 480
Db 4797 CAACGTGGGGGCAAGGACTGGCAGGGGAAGCCAGGCACTTACAGATTGTGGCACCGGGG 4856
Qy 481 GluArgProSerGlyMetPheAspSerSerValLeuCysGluCysTyrAspAlaGlyCys 500
Db 4857 GAGCGCCCTCCGCGCATGTTTCGACTCGTCCGCTCTGTGAGTGTATGACGCGAGGCTGT 4916
Qy 501 AlaTyrTyrGluLeuThrProAlaGluThrThrValArgLeuArgAlaTyrMetAsnThr 520
Db 4917 GCTTGTGTATGAGTCAACGCGCGCGAGACTACAGTTTAGGCTACGAGCGTACATGAACAC 4976
Qy 521 ProGlyLeuProValCysGlnAspHisLeuGluPheThrGluGlyValPheThrGlyLeu 540
Db 4977 CCGGGGCTTCCCGTGTGCCAGGACCATCTTGAATTTTGGAGGGCGCTTTTACAGGCTC 5036
Qy 541 ThrHisIleAspAlaHisPheLeuSerGlnThrLysGlnSerGlyGluAsnLeuProTyr 560
Db 5037 ACTCATATAGATGCCCATCTTCTATCCAGACAAGACAGAGTGGGAGAACCTTCTCTTAC 5096
Qy 561 LeuValAlaTyrGlnAlaThrValCysAlaArgAlaGlnAlaProProProSerThrAsp 580
Db 5097 CTGGTAGCGTACCAAGCCACCGTGTGCGCTAGGGCTCAAGGCCCTCCCCCATCGTGGGAC 5156
Qy 581 GlnMetTyrPlyCysIleuIleArgLeuLysProThrLeuHisGlyProThrProLeuLeu 600
Db 5157 CAGATGTGAAGTGTGTGATTCGCTCAAGCCACCCTCCATGGGCAACACCCCTTGCTA 5216
Qy 601 TyrArgLeuGlyAlaValGlnAsnGluValThrLeuThrHisProValThrLysTyrIle 620
Db 5217 TACAGACTGGCGCTGTTTCCAGATGAATCAACCTGACGACCCAGTCCACCAATACATC 5276
Qy 621 MetThrCysMetSerAlaAspLeuGluValValThrSerThrThrValLeuValGlyGly 640
Db 5277 ATGACATGATGTGGCGCCACCTGGAGGTGCTCACGAGCACCTGGGTGCTCGTTGGCGG 5336
Qy 641 ValLeuAlaAlaLeuAlaAlaTyrCysLeuSerThrGlyCysValValIleValGlyArg 660
Db 5337 GTCTCGCTGCTTTGGCGCGGTATTCCTGTCAACAGGCTGCTGTGTCTATAGTGGCAGG 5396
Qy 661 IleValLeuSerGlyLysProAlaIleIleProAspArgGluValLeuTyrArgGluPhe 680
Db 5397 GTCGTCTTGTCCGGGAAGCGGCAATCATCTGTACAGGGAAGTCTCTTACCGAGAGTTC 5456
Qy 681 AspGluMetGluGluCys 686
Db 5457 GATGAGATGGAAGAGTGC 5474
```

Search completed: February 27, 2005, 04:08:03  
Job time : 946 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: February 27, 2005, 03:44:53 ; Search time 296 Seconds  
(without alignments)  
3792.182 Million cell updates/sec

Title: US-09-930-591-2

Perfect score: 3618

Sequence: 1 MAPITAAQTRGLGCIIT.....PAIIPDRELYREFDEMEEC 686

Scoring table:

BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlp  
-Q=/cgn2\_1/uspto spool p/US0930591/runat\_25022005\_143613\_25700/app query.fasta\_1.839  
-DB=Issued Patents NA -OPMT=fastcap -SUFFIX=rni -MINWATCH=0 -1 -LOOPCT=0  
-LOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi  
-LIST=45 -DOCLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFMT=pco -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US0930591.qcgn 1 105 @runat\_25022005\_143613\_25700 -NCPUS=6 -ICPU=3  
-NO MMAP -LARGQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXI=7

Database :

Issued Patents NA:  
1: /cgn2\_6/ptodata/1/ina/5A.COMB.seq:  
2: /cgn2\_6/ptodata/1/ina/5B.COMB.seq:  
3: /cgn2\_6/ptodata/1/ina/6A.COMB.seq:  
4: /cgn2\_6/ptodata/1/ina/6B.COMB.seq:  
5: /cgn2\_6/ptodata/1/ina/PCFUS.COMB.seq:  
6: /cgn2\_6/ptodata/1/ina/backfileseq1.seq:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3574	98.8	7310	3	US-08-444-818-74
2	3574	98.8	9379	3	Sequence 74, Appli
3	3574	98.8	9379	4	Sequence 1, Appli
4	3574	98.8	9401	1	Sequence 9, Appli
5	3574	98.8	9401	1	Sequence 9, Appli
6	3574	98.8	9401	3	US-08-440-549-9
7	3574	98.8	9401	3	US-08-440-549-9
8	3571	98.7	6785	3	US-08-823-895A-25
9	3571	98.7	8316	3	US-08-444-818-65
10	3570	98.7	9185	3	US-08-444-818-88
11	3570	98.7	9185	3	Sequence 122, App
12	3565	98.5	2058	4	US-08-444-818-123
					Sequence 2, Appli

13	3565	98.5	2058	4	US-09-881-654-1
14	3565	98.5	2058	4	US-10-637-323-1
15	3565	98.5	8987	3	US-08-444-818-137
16	3555	98.3	5360	3	US-08-444-818-53
17	3551	98.1	9646	3	US-08-811-566-1
18	3551	98.1	9646	3	US-09-034-756-1
19	3551	98.1	12980	3	US-08-811-566-5
20	3551	98.1	12980	3	US-09-034-756-5
21	3544	98.0	9599	3	US-09-014-416-2
22	3544	98.0	9599	3	US-09-014-416-6
23	3538	97.8	9401	2	US-08-432-693-1
24	3538	97.8	9416	3	US-08-811-566-19
25	3538	97.8	9416	3	US-08-034-756-19
26	3535	97.7	9379	3	US-08-444-818-176
27	3533	97.7	9365	4	US-09-827-688-7
28	3532	97.6	9401	5	PCT-US91-02225-9
29	3504	96.8	9416	3	US-08-823-895A-26
30	3504	96.8	9416	4	US-10-104-866-13
31	3429	94.8	7989	4	US-09-539-601-10
32	3429	94.8	8001	4	US-09-539-601-7
33	3426	94.7	8637	4	US-09-539-601-4
34	3426	94.7	8638	4	US-10-029-907-7
35	3426	94.7	8638	4	US-10-029-907-24
36	3426	94.7	8638	4	US-10-029-907-25
37	3426	94.7	8639	4	US-10-029-907-1
38	3426	94.7	8643	4	US-10-029-907-4
39	3426	94.7	8649	4	US-09-539-601-13
40	3426	94.7	11076	4	US-09-539-601-1
41	3423	94.6	8001	4	US-09-539-601-22
42	3421	94.6	8638	4	US-10-029-907-6
43	3421	94.6	9413	1	US-09-827-688-6
44	3420	94.5	6039	1	US-08-324-977-11
45	3420	94.5	6039	2	US-08-384-616-11

ALIGNMENTS

RESULT 1

US-08-444-818-74  
; Sequence 74, Application US/08444818  
; Patent No. 6150087  
; GENERAL INFORMATION:  
; APPLICANT: Chien, David Y.  
; APPLICANT: Rutter, William J.  
; TITLE OF INVENTION: NANBV Diagnostics and Vaccines  
; NUMBER OF SEQUENCES: 777  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Chiron Corporation  
; STREET: 4560 Horton Street  
; CITY: Emeryville  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 94608-2916  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA: US/08/444,818  
; FILING DATE:  
; CLASSIFICATION: 424  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/08/403,590  
; FILING DATE: 14-MAR-1995  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Harbin, Alisa A.  
; REGISTRATION NUMBER: 33,895  
; REFERENCE/DOCKET NUMBER: 0110.002  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (508)359-3876  
; TELEFAX: (508)359-3885  
; INFORMATION FOR SEQ ID NO: 74:

SEQUENCE CHARACTERISTICS:  
LENGTH: 7310 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 3..7310  
US-08-444-818-74

## Alignment Scores:

Pred. No.: 0 Length: 7310  
Score: 3574.00 Matches: 672  
Percent Similarity: 99.42% Conservatives: 10  
Best Local Similarity: 97.96% Mismatches: 4  
Query Match: 98.78% Indels: 0  
DB: 3 Gaps: 0

US-09-930-591-2 (1-686) x US-08-444-818-74 (1-7310)

QY	1	MetAlaProIleThrAlaTyrAlaGlnGlnThrArgGlyLeuLeuGlyCysIleIleThr	20
DB	1728	CTGGCGCCATCATCGCGTACGCGCCAGCAGACAAGGGGCTCTTAGGGTGCATAATCAC	1787
QY	21	SerLeuThrGlyArgAspLysAanGlnValGluGlyGluValGlnIleValSerThrAla	40
DB	1788	AGCTTAACCTGGCGCGGACAAAACCAAGTGGAGGGTGGTCCAGATTGTCTCACTGCT	1847
QY	41	AlaGlnThrPheLeuAlaThrCysIleAsnGlyValCysTrpThrValTyrHisGlyAla	60
DB	1848	GCCCAAAACCTTCCTGGCAACGTGCATCAATGGGTGTGTGGACTGTGTACCAACGGG	1907
QY	61	GlyThrArgThrIleAlaSerProLysGlyProValIleGlnMetTyrThrAanValAsp	80
DB	1908	GGAAACGAGGACCATCGCGTCACCAAGGGTCTGTATCCAGATGTATACCAATGTAGAC	1967
QY	81	GlnAspLeuValGlyTrpProAlaProGlnGlyAlaArgSerLeuThrProCysThrCys	100
DB	1968	CAGACCTTGTGGGTGGCCCGCTCCGCAAGGTAGCGGTCAATGACACCTGCATTCG	2027
QY	101	GlySerSerAspLeuTyrLeuValThrArgHisAlaAspValIleProValArgArg	120
DB	2028	GGCTCTCGACCTTTACCTGGTCAAGGACGACGCGATGTCTATCCCGTGGCGCGG	2087
QY	121	GlyAspGlyArgGlySerLeuLeuSerProArgProIleSerTyrLeuLysGlySerSer	140
DB	2088	GGTGATAGCAGGGGACGCTGTCTGCCCGGCCCATTTTCTTACTTGAAGGCTCTCTCG	2147
QY	141	GlyGlyProLeuLeuCysProAlaGlyHisAlaValGlyIlePheArgAlaAlaValCys	160
DB	2148	GGGGTCCGCTGTGTGCCCGGGGACGCGTGGGCATATTTAGGGCGGGGTGTC	2207
QY	161	ThrArgGlyValAlaLysAlaValAspPheIleProValGluSerLeuGluThrThrMet	180
DB	2208	ACCGGTGAGTGGCTAAGGCGGTGACCTTTATCCCTGTGGAGAACCTTAGACACCAATG	2267
QY	181	ArgSerProValPheSerAspAanSerSerProProAlaValProGlnSerTyrGlnVal	200
DB	2268	AGTCCCCCGTGTTCACGGATACTCTCTCCACCAAGTAGTGGCCCAAGCTTCCAGGTG	2327
QY	201	AlaHisLeuHisAlaProThrGlySerGlyLysSerThrLysValProAlaAlaTyrAla	220
DB	2328	GCTACCTCCATGCTCCACAGGCGGCGGCAAAAGCACCAAGGTCCCGGTGCATATGCA	2387
QY	221	AlaGlnGlyTyrLysValLeuValLeuAsnProSerValAlaAlaThrMetGlyPheGly	240
DB	2388	GCTCAGGGCTATAAGGTGTAGTACTCAACCCCTCTGTGTCTGCAACACACTGGGCTT	2447
QY	241	AlaTyrMetSerLysAlaHisGlyIleAspProAsnIleArgThrGlyValArgThrIle	260
DB	2448	GCTTACATGTCCAGGGCTCATGGATCGATCCTTAACATCAGGACCGGGGTGAGAACAT	2507

QY	261	ThrThrGlySerProIleThrTyrSerThrTyrGlyLysPheLeuAlaAspGlyGlyCys	280
DB	2508	ACCACCTGGCAGCCCATCACGTACTCCACTACGCAAGTTCCTTGCCGACGCGGGTGC	2567
QY	281	SerGlyGlyAlaTyrAspIleIleCysAspGluCysHisSerThrThrAspAlaThrSer	300
DB	2568	TCGGGGGGCGCTTATGACATAATAATTTGTGACGAGTGCCTCCACGGATGCCATCC	2627
QY	301	IleLeuGlyIleGlyThrValLeuAspGlnAlaGluThrAlaGlyAlaArgLeuThrVal	320
DB	2628	ATCTTGGGCATCGGCATCTGTCTTGACCAAGCAGAGACTGCGGGGCGAGACTGGTGTG	2687
QY	321	LeuAlaThrAlaThrProProGlySerValThrValProHisProAsnIleGluGluVal	340
DB	2688	CTCGCCACCGCCACCCCTCCGGGCTCCGTCACCTGTGCCCCATCCCAACATCAGAGAGT	2747
QY	341	AlaLeuSerThrThrGlyGluIleProPheTyrGlyLysAlaIleProLeuGluAlaIle	360
DB	2748	GCTCTGTCCACCGGAGAGATCCCTTTTACGGCAAGGCTATCCCTCGAAGTAATC	2807
QY	361	LysGlyGlyArgHisLeuIlePheCysHisSerLysLysCysAspGluLeuAlaAla	380
DB	2808	AAGGGGGGAGACATCTCATCTTCTGTCTATTCAAGAAGAAGTGCAGCACTCGCGCA	2867
QY	381	LysLeuValAlaLeuGlyValAsnAlaValAlaTyrTyrArgGlyLeuAspValSerVal	400
DB	2868	AAGCTGTCCGATTTGGGCATCAATGCGTGGCTACTACCGGGCTCTTGACGTGTCGTC	2927
QY	401	IleProThrSerGlyAspValValValValAlaThrAspAlaLeuMetThrGlyPheThr	420
DB	2928	ATCCCCACCGCGGCGATTTGTGTCGTGGCAACGATGCCCTCATGACCGGCTATACC	2987
QY	421	GlyAspPheAspSerValIleAspCysAanThrCysValThrGlnThrValAspPheSer	440
DB	2988	GGCGACTTCGATCTCGGTGATAGACTGCAATACGTGTGTCCACGACAGCTCGATTCAGC	3047
QY	441	LeuAspProThrPheThrIleGluThrIleThrLeuProGlnAspAlaValSerArgThr	460
DB	3048	CTTGACCTTACCTTACCATTCAGACATCAGCTCCCGCAGGATGCTGTCTCCGCACT	3107
QY	461	GlnArgArgGlyArgThrGlyArgGlyLysProGlyIleTyrArgPheValAlaProGly	480
DB	3108	CAACGTCGGGGGAGGACTGGCAGGGGAAGCAGGCACTTACAGATTGTGGCACCGGG	3167
QY	481	GluArgProSerGlyMetPheAspSerSerValLeuCysGluCysTyrAspAlaGlyCys	500
DB	3168	GAGCGCCCTCTCCGCACTGTTCCGCTCGTCTGTGTGAGTGTCTAGCGAGGCTGT	3227
QY	501	AlaTyrTyrGluLeuThrProAlaGluThrThrValArgLeuArgAlaTyrMetAsnThr	520
DB	3228	GCTTGGTATGAGCTCACGCCCGCGGAGACTACAGTTAGGCTACGAGCGTACATGAACAC	3287
QY	521	ProGlyLeuProValCysGlnAspHisLeuGluPheThrGluGlyValPheThrGlyLeu	540
DB	3288	CCGGGCTTCCCGGTGGCCAGGACCATCTTGAATTTGGAGGGCGCTTTTACAGGCGCTC	3347
QY	541	ThrHisIleAspAlaHisPheLeuSerGlnThrLysGlnSerGlyGluAsnLeuProTyr	560
DB	3348	ACTCATATAGTGCACCTTTCTATCCAGACAAAGCAGAGTGGGAGAACCTTCTCTTAC	3407
QY	561	LeuValAlaTyrGlnAlaThrValCysAlaArgAlaGlnAlaProProSerTrpAsp	580
DB	3408	CTGGTAGCGTACCAGCCCGGTGTGGCTAGGGCTCAAGCCCTCCCCCATCGTGGGAC	3467
QY	581	GlnMetTrpLysCysLeuIleArgLeuLysProThrLeuHisGlyProThrProLeuLeu	600
DB	3468	CAGATGTGAAGGTGTGTGATTCGCCCTCAGCCCAACCCCTCCATGGGCCAACCCCTGCTA	3527
QY	601	TyrArgLeuGlyAlaValGlnAanGluValThrLeuThrHisProValThrLysTyrIle	620
DB	3528	TACAGACTGGCGCTGTTCAGATGAATCAACCTCGCGCACCCAGTCCACCAATACATC	3587
QY	621	MetThrCysMetSerAlaAspLeuGluValValThrSerThrTrpValLeuValGlyGly	640

```
Db 3588 ATGACATGTCATGTCGGCGACCTGGAGTCTGTCACGACACCTGGTCTGTTGGCGGC 3647
Qy 641 ValLeuAlaLeuAlaAlaTyrCysLeuSerThrGlyCysValValIleValcIlyArg 660
Db 3648 GTCTGCTGCTGTTGGCGCGATGCTCTGTCACAGGCTGCTGGTCATAGTGGGCGAG 3707
Qy 661 IleValLeuSerGlyLysProAlaIleIleProAspArgGluValLeuTyrArgGluPhe 680
Db 3708 GTCGCTTGTCCGGGAAGCGGCAATCATACCTGACAGGGAAGTCTCTACCGAGAGTTC 3767
Qy 681 AspGluMetGluGluCys 686
Db 3768 GATGAGATGGAAGAGTGC 3785

RESULT 2
US-09-388-874-1
; Sequence 1, Application US/09388874
; Patent No. 6284249
; GENERAL INFORMATION:
; APPLICANT: Veronique Barban
; TITLE OF INVENTION: VACCINE COMPOSITION FOR PREVENTING OR
; TITLE OF INVENTION: TREATING C HEPATITIS
; FILE REFERENCE: PWC97-03A
; CURRENT APPLICATION NUMBER: US/09/388.874
; CURRENT FILING DATE: 1998-09-02
; EARLIER APPLICATION NUMBER: PCT/FR98/00448
; EARLIER FILING DATE: 1998-03-06
; EARLIER APPLICATION NUMBER: 97/02,887
; EARLIER FILING DATE: 1997-03-06
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 9379
; TYPE: DNA
; ORGANISM: Virus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (320)...(9352)
US-09-388-874-1

Alignment Scores:
Pred. No.: 0 Length: 9379
Score: 3574.00 Matches: 672
Percent Similarity: 99.42% Conservative: 10
Best Local Similarity: 97.96% Mismatches: 4
Query Match: 98.78% Indels: 0
DB: 3 Gaps: 0

US-09-930-591-2 (1-686) x US-09-388-874-1 (1-9379)
Qy 1 MetAlaProIleThrAlaTyrAlaGlnThrArgGlyLeuLeuGlyCysIleIleThr 20
Db 3395 CTGGCGCCCATCAGCGGCTAGCCGACGACGACGAGGCGCCCTCCTAGGCTGATATCA 3454
Qy 21 SerLeuThrGlyArgAspLysAsnGlnValGluGlyGluValGlnIleValSerAla 40
Db 3455 AGCTTAAGTGGCGGGGACAAACCAAGTGGAGGTGAGGTCCAGATTGTGTCACTGCT 3514
Qy 41 AlaGlnThrPheLeuAlaThrCysIleAsnGlyValCysIleThrValTyrHisGlyAla 60
Db 3515 GCCCAAACTTCTCTGGCAAGCTGCATCAATGGGGTGTGCTGACTGTCTACACCGGGCC 3574
Qy 61 GlyThrArgThrIleAlaSerProIleGlyProValIleGlnMetTyrThrAsnValAsp 80
Db 3575 GGAACGAGGACCATCGCGCTCAACCAAGGGTCTCTGTCATCCAGATGTATACCAATGTAG 3634
Qy 81 GlnAspLeuValGlyTyrProAlaProGlnGlyAlaArgSerLeuThrProCysThrCys 100
Db 3635 CAAGACCTTGTGGGTGGCGCTCCGCAAGGTAGCCGCTCATTCACACCCCTGCACTTGC 3694
Qy 101 GlySerSerAspLeuTyrLeuValThrArgHisAlaAspValIleProValArgArg 120
```

```
Db 3695 GGCTCTCTCGGACCTTTACCTGGTGTCAAGAGGACGCGCATGTCTATCCGTCGCGCGCGG 3754
Qy 121 GlyAspGlyArgGlySerLeuLeuSerProArgProIleSerTyrLeuLysGlySerSer 140
Db 3755 GGTGATAGCAGGGGACGCTGCTGTGCCCCCGGCCCATTTCTTACTTTGAAAGGCTCTCTCG 3814
Qy 141 GlyGlyProLeuLeuCysProAlaGlyHisAlaValGlyIlePheArgAlaAlaValCys 160
Db 3815 GGGGGTCTGCTGTTGTGCCCGCGGGACGCGTGGGCATATTTAGGGCGCGCGTGTGC 3874
Qy 161 ThrArgGlyValAlaLysAlaValAspPheIleProValGluSerLeuGluThrThrMet 180
Db 3875 ACCGTGGAGTGGCTAAGCGGTGGACTTTATCCTCTGTGGAGAACCTTAGACAAACATG 3934
Qy 181 ArgSerProValPheSerAspAsnSerSerProAlaValProGlnSerTyrGlnVal 200
Db 3935 AGGTCCCCGGTGTTCACGGATAACTCTCTCCACCACTAGTGCCTCCAGAGCTTCCAGGTG 3994
Qy 201 AlaHisLeuHisAlaProThrGlySerGlyLysSerThrLysValProAlaAlaTyrAla 220
Db 3995 GCTCACCTCCATGCTCCACAGCGCAGCGCAAAAGCACCAAGGTCCCGGCTGCATATGCA 4054
Qy 221 AlaGlnGlyTyrLysValLeuValLeuAsnProSerValAlaAlaThrMetGlyPheGly 240
Db 4055 GCTCAGGGCTATAGGGTGTCTAGTACTCAACCCCTCTGTGTGTGCAACACCTGGGCTTGGT 4114
Qy 241 AlaTyrMetSerLysAlaHisGlyIleAspProAsnIleArgThrGlyValArgThrIle 260
Db 4115 GCTTACATGTCCAAAGGCTCATGGATGATCTTAACATCAGGACCGGGGTGAGAACAT 4174
Qy 261 ThrThrGlySerProIleThrTyrSerThrTyrGlyLysPheLeuAlaAspGlyGlyCys 280
Db 4175 ACCACTGGCAGCCCATCAGTACTCCACCTAGCGCAAGTTCCTTGGCGCGCGGTGC 4234
Qy 281 SerGlyGlyAlaTyrAspIleIleCysAspGluCysHisSerThrAspAlaThrSer 300
Db 4235 TCGGGGGCGCTTATGACATAATAATTTGTGACAGGTGCCACTCCAGGATGCCACATCC 4294
Qy 301 IleLeuGlyIleGlyThrValLeuAspGlnAlaGluThrAlaGlyAlaArgLeuThrVal 320
Db 4295 ATCTGGGCTATGGCACTGCTTGACCAAGCAGAGACTCGGGGGGCGAGACTGGTGTG 4354
Qy 321 LeuAlaThrAlaThrProGlySerValThrValProHisProAsnIleGluGluVal 340
Db 4355 CTGCGCACCGCCACCCCTCGGGCTCGTCACTGTGCCCCCATCCCAACATCGAGGAGTT 4414
Qy 341 AlaLeuSerThrThrGlyGluIleProPheTyrGlyLysAlaIleProLeuGluAlaIle 360
Db 4415 GCTCTGTCCACCGGAGAGATCCCTTTTACGGCAAGGCTATCCCCCTCGAAGTAAATC 4474
Qy 361 LysGlyGlyArgHisLeuIlePheCysHisSerLysLysCysAspGluLeuAlaAla 380
Db 4475 AAGGGGGGAGACATCTCATCTTCTGTCATTCAAGGAAGAGTGCAGAGAACTCGCGCA 4534
Qy 381 LysLeuValAlaLeuGlyValAsnAlaValAlaTyrTyrArgGlyLeuAspValSerVal 400
Db 4535 AAGCTGTGTCATTTGGGCATCAATCGCTGCGCTTACTACCGCGTCTTTGACGTGTCGTC 4594
Qy 401 IleProThrSerGlyAspValValValAlaThrAspAlaLeuMetThrGlyPheThr 420
Db 4595 ATCCGACACCGCGCGATGTTGTGCTGGCAACCGATGCCCTCATGACCGGTATACC 4654
Qy 421 GlyAspPheAspSerValIleAspCysAsnThrCysValThrGlnThrValAspPheSer 440
Db 4655 GGGGACTTCCACTCGGTGATAGACTGCATACGTGTGTACCCAGACAGCTCGATTTTACG 4714
Qy 441 LeuAspProThrPheThrIleGluThrIleThrLeuProGlnAspAlaValSerArgThr 460
Db 4715 CTTGACCTTACCTTCCATTGAGACATCAAGCTCCCCCAGGATGCTGTCTCCCGCACT 4774
Qy 461 GlnArgArgGlyArgThrGlyArgGlyLysProGlyIleTyrArgPheValAlaProGly 480
Db 4775 CAACGTCCGGGCGAGGACTGGCAGGGGAAGCCAGGCATCTACAGATTTGTGGCACCGGGG 4834
```

Qy	481	GluArgProSerGlyMetPheAspSerSerValLeuCysGluCysTyrAspAlaGlyCys	500
Db	4835	GAGCGCCCTCCGGCATGTTTCGACTCGTCCGCTCTGTGAGTGCTATGACGAGGCTGT	4894
Qy	501	AlaTrpTyrGluLeuThrProAlaGluThrThrValArgLeuArgAlaTyrMetAsnThr	520
Db	4995	GCTTGGTATGAGCTACAGCCCGCCGAGACTACAGTTAGGCTACGAGCGGTACATGAACACC	4954
Qy	521	ProGlyLeuProValCysGlnAspHisLeuGluPheTrpGluGlyValPheThrGlyLeu	540
Db	4955	CCGGGGCTTCCCGTGTGTCAGGACCATCTTGAATTTTGGGAGGCGCTCTTTACAGGCCTC	5014
Qy	541	ThrHisIleAspAlaHisPheLeuSerGlnThrLysGlnSerGlyGluAsnLeuProTyr	560
Db	5015	ACTCATATAGATGCCCACTTTCATCCAGACAAAGCAGAGTGGGAGAACCTTCCTTAC	5074
Qy	561	LeuValIaIaTyrGlnAlaThrValCysAlaAargAlaGlnAlaProProProSerTrpAsp	580
Db	5075	CTGCTAGCGTACCAGGACACCGTGTGGCTAGGGCTCAAGCCCTCCCCCATCGTGGGAC	5134
Qy	581	GlnMetTrpLysCysLeuIleArgLeuLysProThrLeuHisGlyProThrProLeuLeu	600
Db	5135	CAGATGTGAAGTGGTTGATTGGCTCAAGCCCACTTCATGGGCCCAACACCCCTGCTA	5194
Qy	601	TyrArgLeuGlyAlaValAlnAsnGluValThrLeuThrHisProValThrLysTyrIle	620
Db	5195	TACAGACTGGCGCTGTTTCAGATGAATATCACCCTGACGACCCAGTCACCAATATCATC	5254
Qy	621	MetThrCysMetSerAlaAspLeuGluValValThrSerThrTrpValLeuValGlyGly	640
Db	5255	ATGACATGCATGTCGGCGGACCTGGAGGTGTCTACGAGCACCTGGGTGCTCGTTGGCGGC	5314
Qy	641	ValLeuAlaIaLeuAlaIaTyrCysLeuSerThrGlyCysValValIleValGlyArg	660
Db	5315	GTCCTGGCTGCTTTGGCGCGGTATTGCTGTCAACAGGCTCGGTGTGCATAGTGGGCAGG	5374
Qy	661	IleValLeuSerGlyLysProAlaIleIleProAspArgGluValLeuTyrArgGluPhe	680
Db	5375	GTCCTCTGTGTCGGGAGCCGGCCNATCATACCTGACAGGAAGTCTCTTACCAGAGTTTC	5434
Qy	681	AspGluMetGluGluCys	686
Db	5435	GATCAGATGGAAGAGTGC	5452

### RESULT 3

```

US-09-916-359-1
; Sequence 1, Application US/09916359
; Patent No. 6538123
; GENERAL INFORMATION:
; APPLICANT: Veronique Barban
; TITLE OF INVENTION: VACCINE COMPOSITION FOR PREVENTING OR
; TITILE OF INVENTION: TREATING C HEPATITIS
; FILE REFERENCE: PMCF97-03A
; CURRENT APPLICATION NUMBER: US/09/916,359
; CURRENT FILING DATE: 2001-07-26
; PRIOR APPLICATION NUMBER: 09/388,874
; PRIOR FILING DATE: 1999-09-02
; PRIOR APPLICATION NUMBER: 97/02,887
; PRIOR FILING DATE: 1997-03-06
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 9379
; TYPE: DNA
; ORGANISM: Virus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (320) ... (9352)
US-09-916-359-1

```

Score:	3574.00	Matches:	672
Percent Similarity:	99.42%	Conservative:	10
Best Local Similarity:	97.96%	Mismatches:	4
Query Match:	98.78%	Indels:	0
DB:	4	Gaps:	0
US-09-930-591-2 (1-686) x US-09-916-359-1 (1-9379)			
QY	1	MetaAlaProIleThrAlaTyrAlaGlnGlnThrArgGlyLeuLeuGlyCysIleIleThr	20
DB	3395	CTGGGGCCCATCAGGGGTACGCCAGCAGCAAGGGGCTCTTAGGGTCATAATCACC	3454
QY	21	SerLeuThrGlyArgAspLysAsnGlnValGluGlyGluValGlnIleValSerThrAla	40
DB	3455	AGCCTAACTGGCCGGGCAAAAACCAAGTGGAGGTCCAGATTGTGTCAACTGCT	3514
QY	41	AlaGlnThrPheLeuAlaThrCysIleAsnGlyValCysTTPThrValTyrHisGlyAla	60
DB	3515	GCCCAAACTTCTGGCAACGTCATCAATGGGGGTGTCTGGACTGTCTACCAACGGGGCC	3574
QY	61	GlyThrArgThrIleAlaSerProLysGlyProValIleGlnMetTyrThrAsnValAsp	80
DB	3575	GGACAGAGGACCATCGCGTCACCAAGGTCTCTGTATCCAGATGATATACCAATGTAGAC	3634
QY	81	GlnAspLeuValGlyTTPProAlaProGlnGlyAlaAlaArgSerLeuThrProCysThrCys	100
DB	3635	CAAGACCTTGTGGGTGGCCCGCTCGCCAAAGGTAGCGCTCAATTGACACCCCTGCATTCG	3694
QY	101	GlySerSerAspLeuTyrLeuValThrArgHisAlaAspValIleProValArgArgArg	120
DB	3695	GGCTCTCGACACTTATACCTGGTGCAGAGCAGCCCGATGTCAATCCGTGCGCCGGCGG	3754
QY	121	GlyAspGlyArgGlySerLeuLeuSerProArgProIleSerTyrLeuLysGlySerSer	140
DB	3755	GGTGATAGCAGGGGACGCTGCTGTGCCCGCCGCCCATTTCTACTTGAAGGCTCTCTCG	3814
QY	141	GlyGlyProLeuLeuCysProAlaGlyHisAlaValGlyIlePheArgAlaAlaValCys	160
DB	3815	GGGGGTCCGCTGTGTGCCCGGGGACCGCGCTGGGCATATTTAGGCGCGGGGTGTGC	3874
QY	161	ThrArgGlyValAlaLysAlaValAspPheIleProValGluSerLeuGluThrThrMet	180
DB	3875	ACCCGTGGAGTGGCTAAGCGGGTGGACTTTATCCCTGTGGAGAACCTTAGAGACAACATG	3934
QY	181	ArgSerProValPheSerAspAsnSerSerProProAlaValProGlnSerTyrGlnVal	200
DB	3935	AGGTCCCGGTGTTACGGATACTCCTCTCCACAGTAGTGGCCCGAGAGCTTCCAGGTG	3994
QY	201	AlaHisLeuHisAlaProThrGlySerGlyLysSerThrLysValProAlaIatyrAla	220
DB	3995	GCTCACTCCATGTCTCCACAGCAGCGGCAAAAGCACCAAGGTCCCGGTGCATATGCA	4054
QY	221	AlaGlnGlyTyrLysValLeuValLeuAsnProSerValAlaAlaThrMetGlyPheGly	240
DB	4055	GCTCAGGGCTATAAGGTGTAGTACTCAACCCCTCTGTTGCTGCAACACTCTGGGCTTGT	4114
QY	241	AlaTyrMetSerLysAlaHisGlyIleAspProAsnIleArgThrGlyValArgThrIle	260
DB	4115	GCTTACATGTCCAAAGGCTCATGGATGTATCTTAACATCAGGACCGGGGTGAGACANT	4174
QY	261	ThrThrGlySerProIleThrTyrSerThrTyrGlyLysPheLeuAlaAspGlyGlyCys	280
DB	4175	ACCACTGGAGCCCATCAGTACTCACCTACGGCAAGTTCCTTCGCCAGCGCGGTGC	4234
QY	281	SerGlyGlyAlaTyrAspIleIleCysAspGluCysHisSerThrAspAlaThrSer	300
DB	4235	TCGGGGGGCGCTTATGACATAATAATTGTGACAGTGGCACTTCCACGGATGCCATCC	4294
QY	301	IleLeuGlyIleGlyThrValLeuAspGlnAlaGluThrAlaGlyAlaArgLeuThrVal	320
DB	4295	ATCTTGGGCATCGGCATGTCTCTTGACCAAGCAGAGACTCGCGGGGCGAGACTGGTTGTG	4354
QY	321	LeuAlaThrAlaThrProProGlySerValThrValProHisProAsnIleGluGluVal	340

Db 4355 CTCGCCACCGCCACCCCTCCGGGCTCCGTCACTGTGCCCCATCCCAACATCGAGAGGTT 4414  
Qy 341 AlaleuSerThrThrGlyGluProPheThrGlyLysAlaIleProLeuGluAlaIle 360  
Db 4415 GCTCTGTCCACCCCGGAGAGATCCCTTTTACGGCAAGGCTATCCCTCGAAGTAATC 4474  
Qy 361 LysGlyArgHisLeuIlePheCysHisSerLysLysCysAspGluLeuAlaAala 380  
Db 4475 AAGGGGGGAGACATCTCATCTCTGTCAATCAAAAGAGAAAGTGCAGGAAGTCCGCGCA 4534  
Qy 381 LysLeuValAlaLeuGlyValAsnAlaValAlaValAlaValAlaValAlaVal 400  
Db 4535 AAGCTGGTCGATGGGATCAATGCGGTGCGCTACTACCGCGGTCTTGAGTGTCCGTC 4594  
Qy 401 IleProThrSerGlyAspValValValAlaValAlaValAlaValAlaValAla 420  
Db 4595 ATCCCGACCGCGGATGTTGTCGTGTGCAACCGATGCGCTCATCGCGGTATACC 4654  
Qy 421 GlyAspPheAspSerValIleAspCysAsnThrCysValThrGlnThrValAspPheSer 440  
Db 4655 GCGCACTTCGACTCGGTGATAGACTGCAATACGTGTGTACCCAGACAGTGCATTTCCAGC 4714  
Qy 441 LeuAspProThrPheThrIleGluThrIleThrLeuProGlnAspAlaValSerArgThr 460  
Db 4715 CTTGACCCCTACCTTCAACATTTGAGCAATACGCTCCCGGAGTGTGTCTCCCGCACT 4774  
Qy 461 GlnArgArgGlyArgThrGlyArgGlyLysProGlyIleTyrArgPheValAlaProGly 480  
Db 4775 CNAAGCTGGGCGAGAGCTGGCAGGGGAGCAGGCACTACAGATTTGTGGCACCGGG 4834  
Qy 481 GluArgProSerGlyMetPheAspSerSerValLeuCysGluCysTyrAspAlaGlyCys 500  
Db 4835 GAGCGCCCTCCGCGATGTTGCGATGCTGCTCTCTGTGAGTGTATGACGCGAGGCTGT 4894  
Qy 501 AlaTrpTyrGluLeuThrProAlaGluThrThrValArgLeuArgAlaTyrMetAsnThr 520  
Db 4895 GCTTGGTATGAGTCACGCGCCGCGAGACTACATTTAGGCTACGAGCGTACATGAACAC 4954  
Qy 521 ProGlyLeuProValCysGlnAspHisLeuGluPheTrpGluGlyValPheThrGlyLeu 540  
Db 4955 CCGGGGCTCCCGTGTGCCAGACATCTTGAATTTGGGAGGGCGCTTTTACAGGCGTC 5014  
Qy 541 ThrHisLeuAspAlaHisPheLeuSerGlnThrLysGlnSerGlyGluAsnLeuProTyr 560  
Db 5015 ACTCATATAGATGCCACTTTCTATCCAGACAAAGCAGAGTGGGAGAACCTTCTCTTAC 5074  
Qy 561 LeuValAlaTyrGlnAlaThrValCysAlaArgAlaGlnAlaProProSerTrpAsp 580  
Db 5075 CTGCTAGCGTACCAAGCCACCGTGTGCGCTAGGCGCTCAAGCGCCCTCCCGCATCGGGAC 5134  
Qy 581 GlnMetTrpLysCysLeuIleArgLeuLysProThrLeuHisGlyProThrProLeuLeu 600  
Db 5135 CAGATGTGGAAAGTGTGATTCGGCTCAAGCCCACTTCAATGGGCAACACCCCTGCTA 5194  
Qy 601 TyrArgLeuGlyAlaValGlnAsnGluValThrLeuThrHisProValThrLysTyrIle 620  
Db 5195 TACAGACTGGGCGCTGTTCAAGTGAATCAACCTCTGAGCCACCCAGTCAACCAATACATC 5254  
Qy 621 MetThrCysMetSerAlaAspLeuGluValValThrThrThrTrpValLeuValGlyGly 640  
Db 5255 ATGACATGCGATGTCCGCGCACTCGAGTGTCTACGAGCACCTGGGTCTGTTGGGCGGC 5314  
Qy 641 ValLeuAlaAlaLeuAlaTyrCysLeuSerThrGlyCysValIleValIleGlyArg 660  
Db 5315 GTCTCTGCTGTTGGCGCGATTTGCTGTCAACAGCGTCTGGTGGTCATAGTGGGCGAG 5374  
Qy 661 IleValLeuSerGlyLysProAlaIleIleProAspArgGluValLeuTyrArgGluPhe 680  
Db 5375 GTCTCTGTTCCGGAGAGCGCGCATCATCTACAGGGAGTCTCTTACCGAGAGTTC 5434  
Qy 681 AspGluMetGluGluCys 686  
|||||

Db 5435 GATGAGATGGAAGAGTGC 5452  
RESULT 4  
US-07-910-760-9  
; Sequence 9, Application US/07910760  
; Patent No. 5683864  
; GENERAL INFORMATION:  
; APPLICANT: Houghton, Michael  
; APPLICANT: Choo, Qui-Lim  
; APPLICANT: Kuo, George  
; TITLE OF INVENTION: Combinations of Hepatitis C virus (HCV)  
; OF INVENTION: Antigens for Use in Immunoassays for Anti-HCV Antibodies  
; NUMBER OF SEQUENCES: 12  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Chiron Corporation  
; STREET: P.O. Box 8097 (Int. Prop. R-440)  
; CITY: Emeryville  
; STATE: CA  
; COUNTRY: U.S.A.  
; ZIP: 94662-8097  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07910,760  
; FILING DATE: 07-JUL-1992  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Blackburn Esq., Robert P.  
; REGISTRATION NUMBER: 30,447  
; REFERENCE/DOCKET NUMBER: 0101.002  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (510) 601-2702  
; TELEFAX: (510) 655-3542  
; INFORMATION FOR SEQ ID NO: 9:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 9401 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: linear  
; MOLECULE TYPE: DNA (genomic)  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: 342..9374  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: 372  
; OTHER INFORMATION: /note= "This amino acid position  
; OTHER INFORMATION: can also be Thr."  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: 867  
; OTHER INFORMATION: /note= "This amino acid position  
; OTHER INFORMATION: can also be Thr."  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: 1341  
; OTHER INFORMATION: /note= "This amino acid position  
; OTHER INFORMATION: can also be Val."  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: 2148  
; OTHER INFORMATION: /note= "This amino acid position  
; OTHER INFORMATION: can also be Ile."  
; FEATURE:  
; NAME/KEY: misc\_feature

```
LOCATION: 2883
OTHER INFORMATION: /note= "This amino acid position
OTHER INFORMATION: can also be Asn."
FEATURE:
NAME/KEY: misc_feature
LOCATION: 3681
OTHER INFORMATION: /note= "This amino acid position
OTHER INFORMATION: can also be Ser."
FEATURE:
NAME/KEY: misc_feature
LOCATION: 3690
OTHER INFORMATION: /note= "This amino acid position
OTHER INFORMATION: can also be Thr."
FEATURE:
NAME/KEY: misc_feature
LOCATION: 4157
OTHER INFORMATION: /note= "This amino acid position
OTHER INFORMATION: can also be Leu."
FEATURE:
NAME/KEY: misc_feature
LOCATION: 4323
OTHER INFORMATION: /note= "This amino acid position
OTHER INFORMATION: can also be Val."
FEATURE:
NAME/KEY: misc_feature
LOCATION: 4701
OTHER INFORMATION: /note= "This amino acid position
OTHER INFORMATION: can also be Tyr."
FEATURE:
NAME/KEY: misc_feature
LOCATION: 4752
OTHER INFORMATION: /note= "This amino acid position
OTHER INFORMATION: can also be Ser."
FEATURE:
NAME/KEY: misc_feature
LOCATION: 5970
OTHER INFORMATION: /note= "This amino acid position
OTHER INFORMATION: can also be Gly."
FEATURE:
NAME/KEY: misc_feature
LOCATION: 6183
OTHER INFORMATION: /note= "This amino acid position
OTHER INFORMATION: can also be His."
FEATURE:
NAME/KEY: misc_feature
LOCATION: 6186
OTHER INFORMATION: /note= "This amino acid position
OTHER INFORMATION: can also be Cys."
FEATURE:
NAME/KEY: misc_feature
LOCATION: 6402
OTHER INFORMATION: /note= "This amino acid position
OTHER INFORMATION: can also be Val."
FEATURE:
NAME/KEY: misc_feature
LOCATION: 7386
OTHER INFORMATION: /note= "This amino acid position
OTHER INFORMATION: can also be Ser."
FEATURE:
NAME/KEY: misc_feature
LOCATION: 7494
OTHER INFORMATION: /note= "This amino acid position
OTHER INFORMATION: can also be Phe."
FEATURE:
NAME/KEY: misc_feature
LOCATION: 7497
OTHER INFORMATION: /note= "This amino acid position
OTHER INFORMATION: can also be Ala."
FEATURE:
NAME/KEY: misc_feature
LOCATION: 7845
OTHER INFORMATION: /note= "This amino acid position
OTHER INFORMATION: can also be Phe."
```

```
FEATURE:
NAME/KEY: misc_feature
LOCATION: 8409
OTHER INFORMATION: /note= "This amino acid position
OTHER INFORMATION: can also be Gly."
FEATURE:
NAME/KEY: misc_feature
LOCATION: 9102
OTHER INFORMATION: /note= "This amino acid position
OTHER INFORMATION: can also be Gly."
FEATURE:
NAME/KEY: misc_feature
LOCATION: 9327
OTHER INFORMATION: /note= "This amino acid position
OTHER INFORMATION: can also be Pro."
US-07-910-760-9
Alignment Scores:
Pred. No.: 0 Length: 9401
Score: 3574.00 Matches: 672
Percent Similarity: 99.42% Conservative: 10
Best Local Similarity: 97.96% Mismatches: 4
Query Match: 98.78% Indels: 0
DB: 1 Gaps: 0
US-09-930-591-2 (1-686) x US-07-910-760-9 (1-9401)
QY 1 MetAlaProIleThrAlaTyrAlaGlnGlnThrArgGlyLeuLeuGlyCysIleIleThr 20
DB 3417 GTGGCGCCCATCACGGCGTACGCCAGACAAAGGGGCTCTTAGGGTGCAATACACC 3476
QY 21 SerLeuThrGlyArgAspLysAsnGlnValGluGlyGluValGlnIleValSerThrAla 40
DB 3477 AGCTTAACCTGGCCGGGACAAAACCAAGTGGAGGGTGAGGTCCAGATTGTGCACTGCT 3536
QY 41 AlaGlnThrPheLeuAlaThrCysIleAsnGlyValCysTrpThrValTyrHisGlyAla 60
DB 3537 GCCCAAACTTCTGTGCAACGTCATCAATGGGGTGTGTGGACTGTCTACCACGGGGCC 3596
QY 61 GlyThrArgThrIleAlaSerProLysGlyProValIleGlnMetTyrThrAsnValAsp 80
DB 3597 GGAACGAGGACCATCGCTGACCAAGGGTCTCTCATCCAGATGATACCAATGATAGAC 3656
QY 81 GlnAspLeuValGlyTrpProAlaProGlnGlyAlaArgSerLeuThrProCysThrCys 100
DB 3657 CAAGACCTTGTGGCTGGCCGCTCCGCAAGGTAGCCGCTCATTCACACCCCTGCTGC 3716
QY 101 GlySerSerAspLeuTyrIleuValThrArgHisAlaAspValIleProValArgArg 120
DB 3717 GGCTCTCGGACCTTTACCTGGTTCAGAGGCACGCCGATGTCAATCCCGTCCGCGCGG 3776
QY 121 GlyAspGlyArgGlySerLeuLeuSerProArgProIleSerTyrIleuLysGlySerSer 140
DB 3777 GGTGATAGAGGGGCGACCTGCTGTCGCCCGGCCCATTTCTCTACTTGAAGAGCTCCTCG 3836
QY 141 GlyGlyProLeuLeuCysProAlaGlyHisAlaValGlyIlePheArgAlaAlaValCys 160
DB 3837 GGGGGTCCGCTGTTGTGCCCGGGGCGACGCGTGGGGCATATTTAGGGCCGGGTGTCG 3896
QY 161 ThrArgGlyValAlaLysAlaValAspPheIleProValGluSerLeuGluThrThrMet 180
DB 3897 ACCCGTGGAGTGGCTAAGCGGTGGACTTTATCCCTGTGTGAGAAACCTAGACACCAATG 3956
QY 181 ArgSerProValPheSerAspAsnSerSerProAlaValProGlnSerTyrGlnVal 200
DB 3957 AGGTCCCCGGTGTTCACGATAACTCTCTCCACAGTGTAGTCCCGAGAGCTTCCAGGTG 4016
QY 201 AlaHisLeuHisAlaProThrGlySerGlyLysSerThrLysValProAlaAlaTyrAla 220
DB 4017 GCTCACTTCATGCTCCACAGGCGGCGCAAAAGCAACCAAGGTCCCGGCTGATATGCA 4076
QY 221 AlaGlnGlyTyrLysValLeuValLeuAsnProSerValAlaAlaThrMetGlyPheGly 240
```

Db 4077 GCTCAGGGCTATAGGTGCTAGTACTCAACCCCTCTGTTGCTGCAACACTGGGCTTTGGT 4136  
Qy 241 AlaTyrMetSerLysAlaHisGlyLeuAspProAsnIleArgThrGlyValArgThrIle 260  
Db 4137 GCTTACATGTCCAAGGCTCATGGGATCATCTTAACATCAGGACCGGGGTGAGAACAAAT 4196  
Qy 261 ThrThrGlySerProIleThrTyrSerThrTyrGlyLysPheLeuAlaAspGlyGlyCys 280  
Db 4197 ACCACTGGCAGCCCATCAGCTATCTCCACCTACCGCAAGTTCTTGGCCGACGGGGGTGC 4256  
Qy 281 SerGlyGlyAlaTyrAspIleIleCysAspGluCysHisSerThrAspAlaThrSer 300  
Db 4257 TCGGGGGGCGCTTATGACATAATTAATTGTGACGAGTCCACTCCAGGATGCCACATCC 4316  
Qy 301 IleLeuGlyIleGlyThrValLeuAspGlnAlaGluThrAlaGlyAlaArgLeuThrVal 320  
Db 4317 ATCTTGGGCATCGGCACATGCTCTTGACCAAGCAGAGACTGGGGGGGCGAGACTGTTGTG 4376  
Qy 321 LeuAlaThrAlaThrProGlySerValThrValProHisProAsnIleGluGluVal 340  
Db 4377 CTCGCCACCGCCACCCCTCGGGGTCTCGTCACTGTGCCCATCCCAACATCGAGAGGTT 4436  
Qy 341 AlaLeuSerThrThrGlyGluIleProPheTyrGlyLysAlaIleProLeuGluAlaIle 360  
Db 4437 GCTCTGTCCACACCGAGAGATCCCTTTTACGGCAAGCTATCCCTCCGAAGTATC 4496  
Qy 361 LysGlyGlyArgHisLeuIlePheCysHisSerLysLysCysAspGluLeuAlaAla 380  
Db 4497 AAGGGGGGAGACATCTCATCTTGTCTATTCAAGAAGAGTCCGACGAACCTCGCCGCA 4556  
Qy 381 LysLeuValAlaLeuGlyValAsnAlaValAlaTyrTyrArgGlyLeuAspValSerVal 400  
Db 4557 AAGCTGTGCGATTGGGCATCAATGCGCGTGGCCCTACTACCGCGTCTTGACGTCCGTC 4616  
Qy 401 IleProThrSerGlyAspValValValAlaThrAspAlaThrMetThrGlyPheThr 420  
Db 4617 ATCCGACACCGCGAGATGTTGTCGTGGCAACCGATGCCCTCATGACCGGTATACC 4676  
Qy 421 GlyAspPheAspSerValIleAspCysAsnThrCysValThrGlnThrValAspPheSer 440  
Db 4677 GCGACTTCGACTCGGTGATAGACTGCATACGTGTGTCAACCCAGACAGTCGATTCAGC 4736  
Qy 441 LeuAspProThrPheThrIleGluThrIleThrLeuProGlnAspAlaValSerArgThr 460  
Db 4737 CTTGACCCCTACCTTCAACATTGAGACATACGCTCCCGAGGATGCTGCTCCCGCACT 4796  
Qy 461 GlnArgArgGlyArgThrGlyArgGlyProGlyIleTyrArgPheValAlaProGly 480  
Db 4797 CAACGTGGGGCAGGACTGGCAGGGGGAAGCCAGGCATCTACAGATTGTGGCACCGGGG 4856  
Qy 481 GluArgProSerGlyMetPheAspSerSerValLeuCysGluCysTyrAspAlaGlyCys 500  
Db 4857 GAGCGCCCTCCGGCATGTTGCACTCGTCCGTCTCTGTGAGTGTATGACGAGGCTGT 4916  
Qy 501 AlaTyrTyrGluLeuThrProAlaGluThrThrValArgLeuArgAlaTyrMetAsnThr 520  
Db 4917 GCTTGGTATAGCTCACGCCCGCGAGACTACAGTTAGGCTACGAGCGTACATGAACACC 4976  
Qy 521 ProGlyLeuProValCysGlnAspHisLeuGluPheTyrGluGlyValPheThrGlyLeu 540  
Db 4977 CCGGGGCTTCCCGTGTCCAGGACCATCTTGAATTTTGGGAGGGCGCTTTTACAGGGCTC 5036  
Qy 541 ThrHisIleAspAlaHisPheLeuSerGlnThrLysGlnSerGlyGluAsnLeuProTyr 560  
Db 5037 ACTCATATAGTCCCACTTCTTATCCAGAACAGAGAGTGGGGAGAACCTTCTTAC 5096  
Qy 561 LeuValAlaTyrGlnAlaThrValCysAlaArgAlaGlnAlaProProSerThrAsp 580  
Db 5097 CTGGTAGCGTACCAGCACCGCTGCGCTAGGCTCAAGCCCTCCCTCCCATCGTGGGAC 5156  
Qy 581 GlnMetTrpLysCysLeuIleArgLeuLysProThrLeuHisGlyProThrProLeuLeu 600  
Db 5157 CAGATGTGAAGTGTGTTGATTGCTCAAGCCCAACCCCTCCATCGGCGCAACACCCCTGCTA 5216

Qy 601 TyrArgLeuGlyAlaValGlnAsnGluValThrLeuThrHisProValThrLysTyrIle 620  
Db 5217 TACAGACTGGCGCGTGTTCAGAAATGAATCACCCTGACGCCAGTCCACCAATACATC 5276  
Qy 621 MetThrCysMetSerAlaAspLeuGluValValThrSerThrTrpValLeuValGlyGly 640  
Db 5277 ATGACATGCATGTCCGGCCGACCTGGAGTGTCTCAGGACACCTGGGTGCTCGTTGGCGGC 5336  
Qy 641 ValLeuAlaAlaLeuAlaTyrCysLeuSerThrGlyCysValValIleValGlyArg 660  
Db 5337 GTCTGGCTGTTGGCGCGTATTGCTGTCAACAGGCTGGGGTTCATAGTGGGAGG 5396  
Qy 661 IleValLeuSerGlyLysProAlaIleIleProAspArgGluValLeuTyrArgGluPhe 680  
Db 5397 GTGCTCTGTCCGGAGCGGCAATCATACCTGACAGGGAAGTCTCTACCGAGAGTTC 5456  
Qy 681 AspGluMetGluGluCys 686  
Db 5457 GATGAGATGGAAGAGTGC 5474  
RESULT 5  
US-08-440-519-9  
; Sequence 9, Application US/08440519  
; Patent No. 5712087  
; GENERAL INFORMATION:  
; APPLICANT: Houghton, Michael  
; APPLICANT: Choo, Qui-Lim  
; APPLICANT: Kuo, George  
; TITLE OF INVENTION: Combinations of Hepatitis C virus (HCV)  
; TITLE OF INVENTION: Antigens for Use in Immunoassays for Anti-HCV Antibodies  
; NUMBER OF SEQUENCES: 12  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Chiron Corporation  
; STREET: P.O. Box 8097 (Int. Prop. R-440)  
; CITY: Emeryville  
; STATE: CA  
; COUNTRY: U.S.A.  
; ZIP: 94662-8097  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/440,519  
; FILING DATE: 12-MAY-1995  
; CLASSIFICATION: 530  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/910,760  
; FILING DATE: 07-JUL-1992  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Blackburn Esq., Robert P.  
; REGISTRATION NUMBER: 30,447  
; REFERENCE/DOCKET NUMBER: 0101.002  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (510) 601-2702  
; TELEFAX: (510) 655-3542  
; INFORMATION FOR SEQ ID NO: 9:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 9401 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: linear  
; MOLECULE TYPE: DNA (genomic)  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: 342..9374  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: 366  
; OTHER INFORMATION: /note= "This amino acid position  
; OTHER INFORMATION: can also be Arg."



```
/
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: 372
/ OTHER INFORMATION: /note= "This amino acid position
/ OTHER INFORMATION: can also be Thr."
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: 867
/ OTHER INFORMATION: /note= "This amino acid position
/ OTHER INFORMATION: can also be Thr."
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: 1341
/ OTHER INFORMATION: /note= "This amino acid position
/ OTHER INFORMATION: can also be Val."
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: 2148
/ OTHER INFORMATION: /note= "This amino acid position
/ OTHER INFORMATION: can also be Ile."
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: 2883
/ OTHER INFORMATION: /note= "This amino acid position
/ OTHER INFORMATION: can also be Asn."
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: 3681
/ OTHER INFORMATION: /note= "This amino acid position
/ OTHER INFORMATION: can also be Ser."
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: 3690
/ OTHER INFORMATION: /note= "This amino acid position
/ OTHER INFORMATION: can also be Thr."
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: 4167
/ OTHER INFORMATION: /note= "This amino acid position
/ OTHER INFORMATION: can also be Leu."
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: 4323
/ OTHER INFORMATION: /note= "This amino acid position
/ OTHER INFORMATION: can also be Val."
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: 4701
/ OTHER INFORMATION: /note= "This amino acid position
/ OTHER INFORMATION: can also be Tyr."
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: 4752
/ OTHER INFORMATION: /note= "This amino acid position
/ OTHER INFORMATION: can also be Ser."
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: 5970
/ OTHER INFORMATION: /note= "This amino acid position
/ OTHER INFORMATION: can also be Gly."
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: 6183
/ OTHER INFORMATION: /note= "This amino acid position
/ OTHER INFORMATION: can also be His."
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: 6186
/ OTHER INFORMATION: /note= "This amino acid position
/ OTHER INFORMATION: can also be Cys."
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: 6402
```

```
/
/ OTHER INFORMATION: /note= "This amino acid position
/ OTHER INFORMATION: can also be Val."
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: 7386
/ OTHER INFORMATION: /note= "This amino acid position
/ OTHER INFORMATION: can also be Ser."
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: 7494
/ OTHER INFORMATION: /note= "This amino acid position
/ OTHER INFORMATION: can also be Phe."
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: 7497
/ OTHER INFORMATION: /note= "This amino acid position
/ OTHER INFORMATION: can also be Ala."
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: 7845
/ OTHER INFORMATION: /note= "This amino acid position
/ OTHER INFORMATION: can also be Phe."
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: 8409
/ OTHER INFORMATION: /note= "This amino acid position
/ OTHER INFORMATION: can also be Gly."
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: 9102
/ OTHER INFORMATION: /note= "This amino acid position
/ OTHER INFORMATION: can also by Gly."
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: 9327
/ OTHER INFORMATION: /note= "This amino acid position
/ OTHER INFORMATION: can also be Pro."
/
/ US-08-440-519-9
/
Alignment Scores:
Pred. No.: 0 Length: 9401
Score: 3574.00 Matches: 672
Percent Similarity: 99.42% Conservative: 10
Best Local Similarity: 97.96% Mismatches: 4
Query Match: 98.78% Indels: 0
DB: 1 Gaps: 0
/
US-09-930-591-2 (1-686) x US-08-440-519-9 (1-9401)
QY 1 MetAlaProIleThrAlaTyrAlaGlnGlnThrArgGlyLeuLeuGlyCysIleIleThr 20
Db 3417 CTGGCGCCCATCAGCGGTACGCCAGCAGACAGAGGGGCTCTTAGGTGCATAATCACC 3476
QY 21 SerLeuThrGlyArgAspLysAsnGlnValGluGlyGluValGlnIleValSerThrAla 40
Db 3477 AGCCTAACTGGCGGACAAAAACCAAGTGGAGGTGAGGTCCAGATTGTGCACTGCT 3536
QY 41 AlaGlnThrPheLeuAlaThrCysIleAsnGlyValCysTrpThrValTyrHisGlyAla 60
Db 3537 GCCCAACCTTCTGTGGCAACGTGCATCAATGGGTGTGTGGACTGTCTACCAACGGGCC 3596
QY 61 GlyThrArgThrIleAlaSerProLysGlyProValIleGlnMetTyrThrAsnValAsp 80
Db 3597 GGNACGAGGACCATCGCGTCACCCAGGGTCTCTGTATCATCATGATATACCAATGTAGAC 3656
QY 81 GlnAspLeuValGlyTrpProLapProGlnGlyAlaArgSerLeuThrProCysThrCys 100
Db 3657 CAAGACCTTGTGGGTGGCGGCTCCGCAAGGTAGCGGCTCATTCAGACCCCTGCACTTGC 3716
QY 101 GlySerSerAspLeuTyrIleuValThrArgHisAlaAspValIleProValArgArg 120
Db 3717 GGCTCTCGGACCTTTTACCTGTGTCAGGACGCGCATGTCTATCCCGTGGCGCGCG 3776
```

QY 121 GlyAspGlyArgGlySerLeuLeuSerProArgProIleSerTyrLeuLysGlySerSer 140  
DB 3777 GGTATACAGAGGGGAGCTGCTGTCGCCCGGCCCATTTCTACTAGAAAGGCTCTCTCG 3836  
QY 141 GlyGlyProLeuLeuCysProAlaGlyHisAlaValAlaGlyIlePheArgAlaAlaValCys 160  
DB 3837 GGGGGTCCGCTGTTGTGCCCCCGGGGACGCCGTGGGCATATTTAGGGCCCGGGTGTGC 3896  
QY 161 ThrArgGlyValAlaLysAlaValAspPheIleProValGluSerLeuGluThrThrMet 180  
DB 3897 ACCCGTGAGTGGCTAAGCGGTGGACTTTATCCCTGTGGAGAACCTTAGAGACAACCATG 3956  
QY 181 ArgSerProValPheSerAspAsnSerSerProProAlaValProGlnSerTyrGlnVal 200  
DB 3957 AGGTCCCCGGTGTTCACGGATAACTCTCTCCACCGAGTAGTCCGCCAGAGCTTCACGGTG 4016  
QY 201 AlaHisLeuHisAlaProThrGlySerGlyLysSerThrLysValProAlaAlaTyrAla 220  
DB 4017 GCTACCTTCATGTCTCCACAGGACGCCGAAAGACCAAGGTCGCCGCTGCATATGCA 4076  
QY 221 AlaGlnGlyTyrLysValLeuValLeuAsnProSerValAlaAlaThrMetGlyPheGly 240  
DB 4077 GCTCAGGGCTATAGGTCTAGTACTCAACCCCTCTGTTGCTGCAACTGGGCTTTGGT 4136  
QY 241 AlaTyrMetSerLysAlaHisGlyIleAspProAsnIleArgThrGlyValArgThrIle 260  
DB 4137 GCTTACATGTCGAAGGCTCATGGGATCGATCCTAACATCAGGACCGGGGTGAGAACAAT 4196  
QY 261 ThrThrGlySerProIleThrTyrSerThrTyrGlyLysPheLeuAlaAspGlyGlyCys 280  
DB 4197 ACCACTGGCAGCCCATCAGCTACTCCACCTACCGCAAGTTCCTTGGCCGACGGGGTGC 4256  
QY 281 SerGlyGlyAlaTyrAspIleIleCysAspGluCysHisSerThrAspAlaThrSer 300  
DB 4257 TCGGGGGCGCTTATGACATATATTTGTGACAGATGCCACTCCACGATGCCACATCC 4316  
QY 301 IleLeuGlyIleGlyThrValLeuAspGlnAlaGluThrAlaGlyAlaArgLeuThrVal 320  
DB 4317 ATCTTGGGCATCGGCACCTGCTTGGACCAAGCAGAGACTGCGGGGGGAGACTGTTGTG 4376  
QY 321 LeuAlaThrAlaThrProGlySerValThrValProHisProAsnIleGluGluVal 340  
DB 4377 CTCCGCCACCGCCATCCCTCGGGGTCCGTGCTACTGTGCCCATCCCAACATCGAGAGGTT 4436  
QY 341 AlaLeuSerThrThrGlyGluIleProPheTyrGlyLysAlaIleProLeuGluAlaIle 360  
DB 4437 GCTCTGTCCACCGGAGAGATCCCTTTTACGGCAAGGCTATCCCTCCGAAAGTATC 4496  
QY 361 LysGlyGlyArgHisLeuIlePheCysHisSerLysLysCysAspGluLeuAlaAla 380  
DB 4497 AAGGGGGGAGACATCTCATCTTGTGCTATTCAAGAAGAGTGGAGCACTCGCCGCA 4556  
QY 381 LysLeuValAlaLeuGlyValAsnAlaValAlaTyrTyrArgGlyLeuAspValSerVal 400  
DB 4557 AAGCTGTGTGCTTGGGCATCAATCGCGCTACTACCGCGTCTTACGCTGTCGCTC 4616  
QY 401 IleProThrSerGlyAspValValValAlaThrAspAlaLeuMetThrGlyPheThr 420  
DB 4617 ATCCCGACCGCGCATGTTGTGCTGTGGCAACCGATGCCCTCATGACCGGTATACC 4676  
QY 421 GlyAspPheAspSerValIleAspCysAsnThrCysValThrGlnThrValAspPheSer 440  
DB 4677 GGCAGCTTCGACTCGGTATAGATCTGCATACGTGTGTCCACCCAGACAGTCGATTTGAGC 4736  
QY 441 LeuAspProThrPheThrIleGluThrIleThrLeuProGlnAspAlaValSerArgThr 460  
DB 4737 CTTCACCTTACCTTACCATTTAGACAATCAGCTCCCGAGGATGCTGTCTCCGCACT 4796  
QY 461 GlnArgArgGlyArgThrGlyArgGlyLysProGlyIleTyrArgPheValAlaProGly 480  
DB 4797 CAACGTGGGCGAGAGCTGGCAGGGGAAGCAGGCATCTACAGATTTGTGGCACCGGG 4856  
QY 481 GluArgProSerGlyMetPheAspSerSerValLeuCysGluCysTyrAspAlaGlyCys 500

DB 4857 GAGCGCCCTCCGCGATGTTCCGACTCGTCCGTCCTCTGTGAGTGTATGACGACAGGCTGT 4916  
QY 501 AlaTyrTyrGluLeuThrProAlaGluThrThrValArgLeuArgAlaTyrMetAsnThr 520  
DB 4917 GCTTGGTATAGCTTACGCCCGCCGAGACTACAGATTAGGCTACGAGCGTACATGAAACAC 4976  
QY 521 ProGlyLeuProValCysGlnAspHisLeuGluPheTyrGluGlyValPheThrGlyLeu 540  
DB 4977 CCGGGGCTTCCCGTGTGCCAGGACCATCTTGAATTTTGGGAGGGCGTCTTTACAGGCTC 5036  
QY 541 ThrHisIleAspAlaHisPheLeuSerGlnThrLysGlnSerGlyGluAsnLeuProTyr 560  
DB 5037 ACTCATATAGATGCCCACTTTCTATCCAGACAAAGCAGAGTGGGGAGAACCTTCTTAC 5096  
QY 561 LeuValAlaTyrGlnAlaThrValCysAlaArgAlaGlnAlaProProProSerTyrAsp 580  
DB 5097 CTGTAGCGTACCAAGCACCGGTGTGCTGAGGCTCAAGCCCTTCCCCCATCGTGGGAC 5156  
QY 581 GlnMetTyrLysCysLeuIleArgLeuLysProThrLeuHisGlyProThrProLeuLeu 600  
DB 5157 CAGATGTGGAAGTGTGTTGATTGCGCTCAAGCCCACTCCATGGGCCAACACCCCTGCTA 5216  
QY 601 TyrArgLeuGlyAlaValGlnAsnGluValThrLeuThrHisProValThrLysTyrIle 620  
DB 5217 TACAGACTGGGCGCTGTTTCAAGTAAATCACCCTGACGCCACCCAGTACCAAAATACATC 5276  
QY 621 MetThrCysMetSerAlaAspLeuGluValValThrSerThrTyrValLeuValGlyGly 640  
DB 5277 ATGACATGCAATGTGCGCGCACCTGAGGTGCTGACGAGCACCTGGGTGCTCGTTGGGGC 5336  
QY 641 ValLeuAlaAlaLeuAlaTyrCysLeuSerThrGlyCysValIleValGlyArg 660  
DB 5337 GTCTGCTGCTTGTGGCGGGTATTGCTGTCAACAGGCTGCGGTGTCATAGTGGGCGG 5396  
QY 661 IleValLeuSerGlyLysProAlaIleIleProAspArgGluValLeuTyrArgGluPhe 680  
DB 5397 GTGCTGTTGTCGGGAAGCGGCAATCATCTGACAGGGAAGTCTCTACCGAGAGTTC 5456  
QY 681 AspGluMetGluGluCys 686  
DB 5457 GATGAGATGGAAGAGTGC 5474  
RESULT 6  
US-08-440-549-9  
; Sequence 9, Application US/08440549  
; Patent No. 6312889  
; GENERAL INFORMATION:  
; APPLICANT: Houghton, Michael  
; APPLICANT: Choo, Qui-Lim  
; APPLICANT: Kuo, George  
; TITLE OF INVENTION: Combinations of Hepatitis C virus (HCV)  
; NUMBER OF SEQUENCES: 12  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Chiron Corporation  
; STREET: P.O. Box 8097 (Int. Prop. R-440)  
; CITY: Emeryville  
; STATE: CA  
; COUNTRY: U.S.A.  
; ZIP: 94662-8097  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/440,549  
; FILING DATE: 12-MAY-1995  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/910,760  
; FILING DATE: 07-JUL-1992



Db 3417 CTGGCGCCCATCAGCGGTACGCCAGACAGACAGGGGCGCTCCTAGGGTGTCATATCACC 3476  
Qy 21 SerLeuThrGlyArgAspGlyAsnGlnValGluGlyGluValGlnIleValSerThrAla 40  
Db 3477 AGCCTAACTGGCCGCGGACAAAACCAAGTGGAGGGTGAGGTCCAGATTGTGTCAACTGCT 3536  
Qy 41 AlaGlnThrPheLeuAlaThrCysIleAsnGlyValCysTyrThrValTyrHisGlyAla 60  
Db 3537 GCCAAACCTTCTTGGCAACGTGCATCAATGGGGTGTCTGGACTGTCTACACAGGGGCC 3596  
Qy 61 GlyThrArgThrIleAlaSerProLysGlyProValIleGlnMetTyrThrAsnValAsp 80  
Db 3597 GGAACGAGACCATCGCGCTCACCCAGAGGTCTCTGTATCCAGATGTATACCAATGTAGAC 3656  
Qy 81 GlnAspLeuValGlyTyrProAlaProGlnGlyAlaArgSerLeuThrProCysThrCys 100  
Db 3657 CAAGACCTTGTGGGTGGCGCGCTCCGCAAGGTAGCCGCTCATTTGACACCCCTGCACCTGC 3716  
Qy 101 GlySerSerAspLeuTyrLeuValThrArgHisAlaAspValIleProValArgArgArg 120  
Db 3717 GGCTCCTCGGACCTTTACCTGGTCAAGAGGACGCCGATGTCAATTCCTGGTGGCGGG 3776  
Qy 121 GlyAspGlyArgGlySerLeuSerProArgProIleSerTyrLeuLysGlySerSer 140  
Db 3777 GGTGATAGCAGGGCAGCGCTGCTGTGCGCCCGGCCCATTTCTACTTGAAGGCTCTCTCG 3836  
Qy 141 GlyGlyProLeuLeuCysProAlaGlyHisAlaValGlyIlePheArgAlaAlaValCys 160  
Db 3837 GGGGGTCCGCTGTTGTGCGCGCGGGCAGCGCGTGGGCATATTTAGGGCGCGGTGTGC 3896  
Qy 161 ThrArgGlyValAlaLysAlaValAspPheIleProValGluSerLeuGluThrThrMet 180  
Db 3897 ACCCGTGGAGTGGCTAAAGCGGTGGACTTTATCTCTGTGGAGAACCTAGAGACCAACATG 3956  
Qy 181 ArgSerProValPheSerAspAsnSerSerProProAlaValProGlnSerTyrGlnVal 200  
Db 3957 AGTCCCGGTGTTTCAAGGATCTCTCTCCACAGTAGTGCCCGAGGCTTCAGGTG 4016  
Qy 201 AlaHisLeuHisAlaProThrGlySerGlyLysSerThrLysValProAlaAlaTyrAla 220  
Db 4017 GCTCACCTCCATGCTCCACAGGCGAGCGGCAAGACCAAGGTCCTCCGCTGCATATGCA 4076  
Qy 221 AlaGlnGlyTyrLysValLeuValLeuAsnProSerValAlaAlaThrMetGlyPheGly 240  
Db 4077 GCTCAGGGCTATAAGGTGTAGTACTCAACCCCTCTGTGTGTCACACACTGGGCTTGTGT 4136  
Qy 241 AlaTyrMetSerLysAlaHisGlyIleAspProAsnIleArgThrGlyValArgThrIle 260  
Db 4137 GCTTACATGTCCAAAGGCTCATGGGATCGATCCTAACATCAGGACCGGGGTGAGAACAT 4196  
Qy 261 ThrThrGlySerProIleThrTyrSerThrTyrLysPheLeuAlaAspGlyGlyCys 280  
Db 4197 ACCACTGGACGCCCATCATCGTACTCCACCTACCGCAAGTTCCTTGGCGAGCGGGTGC 4256  
Qy 281 SerGlyGlyAlaTyrAspIleIleCysAspGluCysHisSerThrAspAlaThrSer 300  
Db 4257 TCGGGGGCGGCTTATGACATAATAATTTGTGACAGTGCACCTCCACGGATGCCACATCC 4316  
Qy 301 IleLeuGlyIleGlyThrValLeuAspGlnAlaGluThrAlaGlyAlaArgLeuThrVal 320  
Db 4317 ATCTTGGGCATCGGCATGCTCTTGGACCAAGCAGAGACTGCGGGGGCGAGACTGGTTGTG 4376  
Qy 321 LeuAlaThrAlaThrProGlySerValThrValProHisProAsnIleGluGluVal 340  
Db 4377 CTGCGCACCGCCACCCCTCGGGGTCTGTCATGTGCCCCATCCCAACATCGAGAGAGTT 4436  
Qy 341 AlaLeuSerThrThrGlyGluIleProPheTyrGlyLysAlaIleProLeuGluAlaIle 360  
Db 4437 GCTCTGTCCACCGGAGAGATCCCTTTTACGGCAAGGCTATCCCTCCCTCGAAGTAATC 4496  
Qy 361 LysGlyGlyArgHisLeuIlePheCysHisSerLysLysCysAspGluLeuAlaAla 380  
Db 4497 AAGGGGGGAGACATCTCATCTTGTCAATTCAAAGAGAGGTGCGAGCAACTCGCGCA 4556

Qy 381 LysLeuValAlaLeuGlyValAsnAlaValAlaTyrTyrArgGlyLeuAspValSerVal 400  
Db 4557 AAGTGTGTGTCATTTGGGCATCAATGCGGTGCTTACTACCGCGTCTTTGACGTGTCCGTC 4616  
Qy 401 IleProThrSerGlyAspValValValAlaThrAspAlaLeuMetThrGlyPheThr 420  
Db 4617 ATCCCGACCAAGCGCGATTTGTGTCGTGGCAACCGATGCCCTCATGACCGGTATACC 4676  
Qy 421 GlyAspPheAspSerValIleAspCysAsnThrCysValThrGlnThrValAspPheSer 440  
Db 4677 GGCACCTTCGACTCGGTGATAGACTGCATATACGTGTGTACCCAGACAGTGCATTTGAC 4736  
Qy 441 LeuAspProThrPheThrIleGluThrIleThrLeuProGlnAspAlaValSerArgThr 460  
Db 4737 CTTGACCTTACCTTACCAATGAGACAATCAGCTCCCCAGGATGTGTCTCCCGCACT 4796  
Qy 461 GlnArgArgGlyArgThrGlyArgGlyLysProGlyIleTyrArgPheValAlaProGly 480  
Db 4797 CAACGTGGGGCAGGACTGGCAGGGGAAGCAGGACATCTACAGATTTGTGGCACCGGG 4856  
Qy 481 GluArgProSerGlyMetPheAspSerSerValLeuCysGluCysTyrAspAlaGlyCys 500  
Db 4857 GAGCGCCCTTCCGCGATTTGACTCGTCTGCTCTCTGTGAGTCTATGACGAGGCTGT 4916  
Qy 501 AlaTyrTyrGluLeuThrProAlaGluThrThrValArgLeuArgAlaTyrMetAsnThr 520  
Db 4917 GCTTGGTATGAGTCAACGCCCGCGAGACTACAGTTAGGCTACGAGGTGATCAACAC 4976  
Qy 521 ProGlyLeuProValCysGlnAspHisLeuGluPheTyrGluGlyValPheThrGlyLeu 540  
Db 4977 CCGGGGCTTCCGCTGTGCCAGGACCATTTGAATTTGGGAGGGCGCTTTTACAGGCTC 5036  
Qy 541 ThrHisIleAspAlaHisPheLeuSerGlnThrLysGlnSerGlyGluAsnLeuProTyr 560  
Db 5037 ACTCATATAGTCCCACTTTCATCCAGAAAGCAGAGTGGGAGAACCTTCTTCTTAC 5096  
Qy 561 LeuValAlaTyrGlnAlaThrValCysAlaArgAlaGlnAlaProProSerTyrAsp 580  
Db 5097 CTGGTAGGTACCAAGCCACCGTGTGCGCTAGGGCTCAAGCCCTCCCCCATCGTGGGAC 5156  
Qy 581 GlnMetTyrLysCysLeuIleArgLeuLysProThrLeuHisGlyProThrProLeuLeu 600  
Db 5157 CAGATGGGAAGTGTGATTCGCTCAAGCCACCTCCATGGGCCCAACACCCCTGCTA 5216  
Qy 601 TyrArgLeuGlyAlaValGlnAsnGluValThrLeuThrHisProValThrLysTyrIle 620  
Db 5217 TACAGACTGGGGCGTGTTCAGAAATGAATCACCTGACGCCACCCAGTCACCAATATC 5276  
Qy 621 MetThrCysMetSerAlaAspLeuGluValValThrSerThrTyrValLeuValGlyGly 640  
Db 5277 ATGACATGCATGTCCGCGACCTGGAGTGTCTACGAGCACCTGGGTGCTCGTGGCGGC 5336  
Qy 641 ValLeuAlaAlaLeuAlaTyrCysLeuSerThrGlyCysValValIleValGlyArg 660  
Db 5337 GTCTGTGCTCTTGGCGGGTATTCGCTCAACAGGCTGCGTGGTCATAGTGGGCGAG 5396  
Qy 661 IleValLeuSerGlyLysProAlaIleIleProAspArgGluValLeuTyrArgGluPhe 680  
Db 5397 GTGCTCTTGTCCGGGAAGCGGCAATCATACCTGACAGGGAAGTCTCTTACCGAGAGTTC 5456  
Qy 681 AspGluMetGluGluCys 686  
Db 5457 GATGAGATGGAGAGTGC 5474

## RESULT 7

US-08-823-895A-25

; Sequence 25, Application US/08823895A

; Patent No. 6433159

; GENERAL INFORMATION:

; APPLICANT: Kevin P. Anderson

; TITLE OF INVENTION: Compositions And Methods For

; TITLE OF INVENTION: Treatment Of Hepatitis C Virus-Associated Diseases

/ NUMBER OF SEQUENCES: 27  
/ CORRESPONDENCE ADDRESS:  
/ ADDRESSEE: Jane Massey Licata, Esq.  
/ STREET: 66 E. Main Street  
/ CITY: Marlton  
/ STATE: NJ  
/ COUNTRY: USA  
/ ZIP: 08053  
/ COMPUTER READABLE FORM:  
/ MEDIUM TYPE: DISKETTE, 3.5 INCH, 1.44 Mb STORAGE  
/ COMPUTER: IBM 486  
/ OPERATING SYSTEM: WINDOWS FOR WORKGROUPS  
/ SOFTWARE: WORDPERFECT 5.1  
/ CURRENT APPLICATION DATA:  
/ APPLICATION NUMBER: US/08/823,895A  
/ FILING DATE: March 17, 1997  
/ CLASSIFICATION: 514  
/ PRIOR APPLICATION DATA:  
/ APPLICATION NUMBER: 08/453,085  
/ FILING DATE: May 30, 1995  
/ PRIOR APPLICATION DATA:  
/ APPLICATION NUMBER: 07/945,289  
/ FILING DATE: September 10, 1992  
/ ATTORNEY/AGENT INFORMATION:  
/ NAME: Jane Massey Licata  
/ REGISTRATION NUMBER: 32,257  
/ REFERENCE/DOCKET NUMBER: ISPH-0203  
/ TELECOMMUNICATION INFORMATION:  
/ TELEPHONE: (609) 779-2400  
/ TELEFAX: (609) 810-1454  
/ INFORMATION FOR SEQ ID NO: 25:  
/ SEQUENCE CHARACTERISTICS:  
/ LENGTH: 9401  
/ TYPE: Nucleic  
/ STRANDEDNESS: Single  
/ TOPOLOGY: Linear  
/ ANTI-SENSE: NO  
/ US-08-823-895A-25

Alignment Scores:  
Pred. No.: 0 Length: 9401  
Score: 3574.00 Matches: 672  
Percent Similarity: 99.42% Conservative: 10  
Best Local Similarity: 97.96% Mismatches: 4  
Query Match: 98.78% Indels: 0  
Gaps: 0  
DB: 3

US-09-930-591-2 (1-686) x US-08-823-895A-25 (1-9401)

QY 1 MetAlaProIleThrAlaTyrAlaGlnGlnThrArgGlyLeuLeuGlyCysIleIleThr 20  
DB 3417 CTGGCGCCCATCAGCGGTACGCCAGCAGACAAAGGGGCTCTAGGTGCATAATCACC 3476  
QY 21 SerLeuThrGlyArgAspLysAsnGlnValGluGlyValGlnIleValSerThrAla 40  
DB 3477 AGCCTAACTGGCGGGACAAAACCAAGTGGAGGTGAGGTCCAGATTGTGCTCAACTGCT 3536  
QY 41 AlaGlnThrPheLeuAlaThrCysIleAsnGlyValCysTrpThrValTyrHisGlyAla 60  
DB 3537 GCCCAACCTCTCTGGCAACGTGCATCAATGGGGGTGCTGGAGCTGTCTACCAAGGGGCC 3596  
QY 61 GlyThrArgThrIleAlaSerProLysGlyProValIleGlnMetTyrThrAsnValAsp 80  
DB 3597 GGAACGAGGACCATCGGCTCAACCAAGGGTCTCTATCCAGATGATATACCAATGTAGAC 3656  
QY 81 GlnAspLeuValGlyTrpProAlaProGlnGlyAlaArgSerLeuThrProCysThrCys 100  
DB 3657 CAAGACCTTGTGGGCTGGCGCGCTCCGCAAGGTAGCGCTCAATTGACACCCCTGCACCTGC 3716  
QY 101 GlySerSerAspLeuThrValThrArgHisAlaAspValIleProValArgArgArg 120  
DB 3717 GGCTCTCTGGACCTTTTACCTTGGTCCAGGACGCCGATGTCTATTTCCGTCGCCCGCGG 3776

QY 121 GlyAspGlyArgGlySerLeuLeuSerProArgProIleSerTyrIleuLysGlySerSer 140  
DB 3777 GGTGATAGCAGGGGCGAGCCCTGCTGTCGCCCGGCCCATTTCTTACTTTGAAAGGCTCCTCG 3836  
QY 141 GlyGlyProLeuLeuCysProAlaGlyHisAlaValGlyIlePheArgAlaAlaValCys 160  
DB 3837 GGGGGTCCGCTGTGTGCCCCCGGGGCGACCGCGGGGCATATTTAGGGCCGGGGTGTGC 3896  
QY 161 ThrArgGlyValAlaLysAlaValAspPheIleProValIleSerLeuGluThrThrMet 180  
DB 3897 ACCGTGGAGTGGCTAAGCGGTGGACTTTATCCCTGTGGAGAACCTAGACACCAATG 3956  
QY 181 ArgSerProValPheSerAspAsnSerSerProAlaValProGlnSerTyrGlnVal 200  
DB 3957 AGGTCCCGGTGTTCACGGATAACTCTCTCCACAGTAGTGCCTCCAGAGCTTCCAGGTG 4016  
QY 201 AlaHisLeuHisAlaProThrGlySerGlyLysSerThrLysValProAlaAlaTyrAla 220  
DB 4017 GGTCACTCCATGCTCCACAGCGAGCGGCAAAAGCACCAAGGTCCCGCTGCATATGCA 4076  
QY 221 AlaGlnGlyTyrLysValLeuValLeuAsnProSerValAlaAlaThrMetGlyPheGly 240  
DB 4077 GCTCAGGGCTATAGGTGTAGTACTCAACCCCTCTGTGCTGCAACACTGGGCTTGGT 4136  
QY 241 AlaTyrMetSerLysAlaHisGlyIleAspProAsnIleArgThrGlyValArgThrIle 260  
DB 4137 GCTTACATGTCCAAGGCTCATGGATCGATCTTAACATCAGGACCGGGGTGAGAACAAAT 4196  
QY 261 ThrThrGlySerProIleThrTyrSerThrTyrGlyLysPheLeuAlaAspGlyLysCys 280  
DB 4197 ACCACTGGCAGGCCCATCAGTACTCCACTACGGCAAGTCTCTGCCAGCGCGGTGTC 4256  
QY 281 SerGlyGlyAlaTyrAspIleIleLysCysAspGluCysHisSerThrAspAlaThrSer 300  
DB 4257 TCGGGGGGCGCTTATGACATATAATTTGTGACGAGTGCCTCCACGSGATGCCACATCC 4316  
QY 301 IleLeuGlyIleGlyThrValLeuAspGlnAlaGluThrAlaGlyAlaArgLeuThrVal 320  
DB 4317 ATCTTGGGCATCGGCATCTGCTTACCAAGCAGAGACTCGCGGGGCGAGACTGGTGTGTG 4376  
QY 321 LeuAlaThrAlaThrProGlySerValThrValProHisProAsnIleGluGluVal 340  
DB 4377 CTCGCCACCGCCACCCCTCCGGGCTCCGTCACCTGTCGCCCATCCCCAACATCGAGGAGTT 4436  
QY 341 AlaLeuSerThrThrGlyGluIleProPheTyrGlyLysAlaIleProLeuGluAlaIle 360  
DB 4437 GCTCTGTCCACCACCGAGAGATCCCTTTTACGGCAAGGCTATCCCTCCGAGTAATC 4496  
QY 361 LysGlyGlyArgHisLeuIlePheCysHisSerLysLysCysAspGluLeuAlaAla 380  
DB 4497 AAGGGGGGAGACATCTCATCTTCTGTCAATCAAAGAGAAGTGCAGAACCTCGCGCA 4556  
QY 381 LysLeuValAlaLeuGlyValAlaValAlaTyrTyrArgGlyLeuAspValSerVal 400  
DB 4557 AGCTGTCGATTTGGGCATCAATCGGTGGCTCTACTACCGCGTCTTGACGTTCGTC 4616  
QY 401 IleProThrSerGlyAspValValValAlaThrAspAlaLeuMetThrGlyPheThr 420  
DB 4617 ATCCCGACCGCGGCGATGTTGTCGTGGTGCACCGATGCCCTCATGCCGCTATACC 4676  
QY 421 GlyAspPheAspSerValIleAspCysAsnThrCysValThrGlnThrValAspPheSer 440  
DB 4677 GGCAGCTTCGACTCGGTGATAGACTGCAATACGTGTGTCCACGACAGTCGATTTTCAGC 4736  
QY 441 LeuAspProThrPheThrIleGluThrIleThrLeuProGlnAspAlaValSerArgThr 460  
DB 4737 CTTGACCCCTTACCTTCCACATGAGACATACCGCTCCCCCAGGATGCTGTCTCCGCACT 4796  
QY 461 GlnArgArgGlyArgThrGlyArgGlyLysProGlyIleTyrArgPheValAlaProGly 480  
DB 4797 CRAAGTCGGGGCAGGACTGGCAGGGGGAAGCCAGGATCTACAGATTTGTGGCACCGGGG 4856  
QY 481 GluArgProSerGlyMetPheAspSerSerValLeuCysGluCysTyrAspAlaGlyCys 500

Db 4857 GAGCCGCTCCGGGATGTTGACCTCGTCCGTCCTCTGTGAGTGTATGACGAGCGTGT 4916  
Qy 501 AlaTptYrGluLeuThrProAlaGluThrThrValArgLeuArgAlaTyrMetAsnThr 520  
Db 4917 GCTTGGTATGAGCTCAGCCCGCCGAGACTACAGTTAGGCTACGAGCGTACATGAACACC 4976  
Qy 521 ProGlyLeuProValCysGlnAspHisLeuGluPheThrGluGlyValPheThrGlyLeu 540  
Db 4977 CGGGGGCTTCGCGTGTGCAGACCATCTTGAATTTTGGGAGGGCGTCTTTACAGGCGTC 5036  
Qy 541 ThrHisLeuAspAlaHisPheLeuSerGlnThrLysGlnSerGlyGluAsnLeuProTyr 560  
Db 5037 ACTCATATAGATGCCACTTCTATCCAGACAAAGCAGAGTGGGGAGAACTTCTCTTAC 5096  
Qy 561 LeuValAlaTyrGlnAlaThrValCysAlaArgAlaGlnAlaProProSerTrpAsp 580  
Db 5097 CTGTGTAGCTACCAAGCCACCGTGTGCTAGGCGTCAAGCCCTCCCTCCCATCGTGGAC 5156  
Qy 581 GlnMetTrpLysCysLeuLeuArgLeuLysProThrLeuHisGlyProThrProLeuLeu 600  
Db 5157 CAGATGTGGAAGTGTGTGATTCGCTCAAGCCACCTCCATGGGCCAACACCCCTGCTA 5216  
Qy 601 TyrArgLeuGlyAlaValGlnAsnGluValThrLeuThrHisProValThrLysTyrIle 620  
Db 5217 TACAGACTGGGGCGCTGTTTCAATGAATCAACCTGACGCCACCTGACCAAAATACATC 5276  
Qy 621 MetThrCysMetSerAlaAspLeuGluValValThrSerThrTrpValLeuValGlyGly 640  
Db 5277 ATGACATGATGTCGGCCGACCTGGAGTGTGTCAGGACACTGGTGTCTGTTGGCGGC 5336  
Qy 641 ValLeuAlaAlaLeuAlaAlaTyrCysLeuSerThrGlyCysValValIleValGlyArg 660  
Db 5337 GTCTGTGCTGTTTGGCGCGCTATTGCTGTCAACAGGCTGCTGATAGTGGGCAGG 5396  
Qy 661 IleValLeuSerGlyLysProAlaIleIleProAspArgGluValLeuTyrArgGluPhe 680  
Db 5397 GTCTGTGCTGTTTGGCGCGCTATTGCTGTCAACAGGCTGCTGATAGTGGGCAGG 5456  
Qy 681 AspGluMetGluGluCys 686  
Db 5457 GATGAGTGGAGAGTGC 5474

## RESULT 8

US-08-444-818-65  
; Sequence 65, Application US/08444818  
; Patent No. 6150087

GENERAL INFORMATION:  
; APPLICANT: Chien, David Y.  
; APPLICANT: Rutter, William J.  
; TITLE OF INVENTION: NANBV Diagnostics and Vaccines  
; NUMBER OF SEQUENCES: 777  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Chiron Corporation  
; STREET: 4560 Horton Street  
; CITY: Emeryville  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 94608-2916  
COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA: US/08/444, 818  
; FILING DATE:  
; CLASSIFICATION: 424  
; PRIOR APPLICATION DATA: US/08/403, 590  
; APPLICATION NUMBER: 14-WAR-1995  
; FILING DATE: 14-WAR-1995  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Harbin, Alisa A.

REGISTRATION NUMBER: 33,895  
REFERENCE/DOCKET NUMBER: 0110.002  
TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (508)359-3876  
; TELEFAX: (508)359-3885  
; INFORMATION FOR SEQ ID NO: 65:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 6785 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: cDNA  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: 3..6785  
US-08-444-818-65

## Alignment Scores:

Pred. No.: 0 Length: 6785  
Score: 3571.00 Matches: 671  
Percent Similarity: 99.42% Conservative: 11  
Best Local Similarity: 97.81% Mismatches: 4  
Query Match: 98.70% Indels: 0  
DB: 3 Gaps: 0

US-09-930-591-2 (1-686) x US-08-444-818-65 (1-6785)

Qy 1 MetAlaProIleThrAlaTyrAlaGlnGlnThrArgGlyLeuLeuGlyCysIleIleThr 20  
Db 1203 CTGGCGCCCATCAGCGGTGTAGCCCGACAGACAGGGGCTCTTAGGGTGCATATATCACC 1262  
Qy 21 SerLeuThrGlyArgAspLysAsnGlnValGluGlyGluValGlnIleValSerThrAla 40  
Db 1263 AGCTAACTGCGCGGACAAAACCAAGTGAGGGTGAGGTCCAGATTGTGTCAACTGCT 1322  
Qy 41 AlaGlnThrPheLeuAlaThrCysIleAsnGlyValCysTrpThrValTyrHisGlyAla 60  
Db 1323 GCCCAAACTTCTCTGGCAACGTCATCAATGGGGTGTGCTGGACTGTCTACACGGGGCC 1382  
Qy 61 GlyThrArgThrIleAlaSerProLysGlyProValIleGlnMetTyrThrAsnValAsp 80  
Db 1383 GGAACGAGAGACCATCGGTACCCAGGGTCTCTCATCCAGATGTATACCAATGTAGAC 1442  
Qy 81 GlnAspLeuValGlyTyrProAlaProGlnGlyAlaArgSerLeuThrProCysThrCys 100  
Db 1443 CAGACCTTGTGGCTGCGCCGCTCCGCAAGGTAGCGCTCATTTGACACCTCGACATTGC 1502  
Qy 101 GlySerSerAspLeuTyrLeuValThrArgHisAlaAspValIleProValArgArgArg 120  
Db 1503 GGCTCTCTCGGACCTTACCTGGTCACGAGGCACGCCGATGTCTATCCCGTGCAGCGCGG 1562  
Qy 121 GlyAspGlyArgGlySerLeuLeuSerProArgProIleSerTyrLeuLysGlySerSer 140  
Db 1563 GGTGATAGCAGGGGCGACCTGCTGTCGCCCGGCCCATTTCTTACTTGAAGGGTCTCCG 1622  
Qy 141 GlyGlyProLeuLeuCysProAlaGlyHisAlaValGlyIlePheArgAlaAlaValCys 160  
Db 1623 GGGGGTCCGCTGTGTGTCGCCCGGGGCAAGCGGTGGGCATATTTAGGGCCGCGGTGTC 1682  
Qy 161 ThrArgGlyValAlaLysAlaValAspPheIleProValGluSerLeuGluThrThrMet 180  
Db 1683 ACCCGTGGAGTGGCTAAGCGGTGACCTTTATCTCTGAGAGAACCTTAGACAAACCATG 1742  
Qy 181 ArgSerProValPheSerAspAsnSerSerProAlaValProGlnSerTyrGlnVal 200  
Db 1743 AGGTCCCGGTGTTTACCGGATAACTCTCTCCACAGTAGTGGCCCGACAGCTCCAGGTG 1802  
Qy 201 AlahisLeuHisAlaProThrGlySerGlyLysSerThrLysValProAlaAlaTyrAla 220  
Db 1803 GCTCACCTCTCATGCTCCACAGGCGGCAAGACACCGAGGTCCCGGCTGCATATGCA 1862  
Qy 221 AlaGlnGlyTyrLysValLeuValLeuAsnProSerValAlaAlaThrMetGlyPheGly 240

Db 1863 GCTCAGGCTATAAGTGTAGTACTCAACCCCTCTGTTGCTGCAACACTGGGCTTTGGT 1922  
Qy 241 AlaTyrMetSerLysAlaHisGlyIleAspProAsnIleArgThrGlyValArgThrIle 260  
Db 1923 GCTTACATGTCGAAGGCTCATGGATCGATCCTTAACATCAGGACCCGGGTGAGAACAT 1982  
Qy 261 ThrThrGlySerProIleThrTyrSerThrTyrGlyLysPheLeuAlaAspGlyGlyCys 280  
Db 1983 ACCACTGGCAGCCCATCATGTAATCCACCTACGCAAGTTCCTTCCCGACGCGGGTGC 2042  
Qy 281 SerGlyGlyAlaTyrAspIleIleCysAspGluCysHisSerThrAspAlaThrSer 300  
Db 2043 TCGGGGGGCGCTTATGACATAATAATTTGTGACCATGTCACCATCCACGATGCGACATCC 2102  
Qy 301 IleLeuGlyIleGlyThrValLeuAspGlnAlaGluThrAlaGlyAlaArgLeuThrVal 320  
Db 2103 ATCTTGGGCATCGGCATGCTTCCTTGACCAAGCAGAGACTCGGGGGGCGAGACTGGTTTG 2162  
Qy 321 LeuAlaThrAlaThrProGlySerValThrValProHisProAsnIleGluGluVal 340  
Db 2163 CTCCGACCGCCACCCCTCGGGCTCGTCACTGTCGCCCATCCCAACATCGAGGAGTT 2222  
Qy 341 AlaLeuSerThrThrGlyGluIleProPheTyrGlyLysAlaIleProLeuGluAlaIle 360  
Db 2223 GCTCTGTCCACCCGAGAGATCCCTTTTACGGCAAGGCTATCCCTCTCGAAGTAATC 2282  
Qy 361 LysGlyGlyArgHisIleIlePheCysHisSerLysLysCysAspGluLeuAlaAla 380  
Db 2283 AAGGGGGGAGACATCTCATCTTCTGTCAATCAAGAAGAGTGCAGCAACTCGCCGCA 2342  
Qy 381 LysLeuValAlaLeuGlyValAsnAlaValAlaTyrTyrArgGlyLeuAspValSerVal 400  
Db 2343 AAGCTGGTCGATGGGCATCATCGTGGTCTACTACCGCGTCTTGACGTCGTC 2402  
Qy 401 IleProThrSerGlyAspValValValAlaThrAspAlaLeuMetThrGlyPheThr 420  
Db 2403 ATCCGACCAACGCGCATGTGTCTGTCGTGGCAACCGATGCCCTCATGACCGCTATACC 2462  
Qy 421 GlyAspPheAspSerValIleAspCysAsnThrCysValThrGlnThrValAspPheSer 440  
Db 2463 GGGGACTTCGACTCGGAGTAGACTGCAATACGTGTGTACCCAGACAGCTCGAATTCAGC 2522  
Qy 441 LeuAspProThrPheThrIleGluThrIleThrLeuProGlnAspAlaValSerArgThr 460  
Db 2523 CTTGACCTTACCTTCCATTTAGACATCATCAGCTCCCCAGAGTGTCTCTCCGACT 2582  
Qy 461 GlnArgGlyArgThrGlyArgGlyLysProGlyIleTyrArgPheValAlaProGly 480  
Db 2583 CAACGTCGGGGCAGGACTGCGAGGGGAGCCAGGCATCTACAGATTTGTGGCACCGGGG 2642  
Qy 481 GluArgProSerGlyMetPheAspSerSerValLeuCysGluCysTyrAspAlaGlyCys 500  
Db 2643 GAGCGCCCTCCGGCATGTTGCACTCGTCTGCTGTGAGTGTGTATGACGAGGCTGT 2702  
Qy 501 AlaTyrThrGluLeuThrProAlaGluThrThrValArgLeuArgAlaTyrMetAsnThr 520  
Db 2703 GCTTGTATGAGTCACGCCCGCGAGACTACAGTTAGGTACGAGGTACATGACACC 2762  
Qy 521 ProGlyLeuProValCysGlnAspHisLeuGluPheThrGluGlyValPheThrGlyLeu 540  
Db 2763 CCGGGGCTTCCTCGTGTCCAGGACCATCTTGAATTTTGGGAGGCGCTTTTACAGGCTC 2822  
Qy 541 ThrHisIleAspAlaHisPheLeuSerGlnThrLysGlnSerGlyGluAsnLeuProTyr 560  
Db 2823 ACTCATATGATGCCACTTTCATCCAGAACAGAGTGGGGAGAACCTTCTCTTAC 2882  
Qy 561 LeuValAlaTyrGlnAlaThrValCysAlaArgAlaGlnAlaProProSerThrAsp 580  
Db 2883 CTGGTAGCTACCAAGCACCGTGTGGCTAGGGCTCAAGCCCTTCCCATCGTGGAC 2942  
Qy 581 GlnMetTrpLysCysLeuIleArgLeuLysProThrLeuHisGlyProThrProLeuLeu 600  
Db 2943 CAGATGTGGAAGTGTGTTGATTCGCTCAAGCCACCCCTCCATGGGCCCAACACCCCTGCTA 3002

Qy 601 TyrArgLeuGlyAlaValGlnAsnGluValThrLeuThrHisProValThrLysTyrIle 620  
Db 3003 TACAGACTGGCGCTGTTTCAAGATGAATCAACCTGACGACCCAGTCCACCAATACATC 3062  
Qy 621 MetThrCysMetSerAlaAspLeuGluValValThrSerThrTrpValLeuValGlyGly 640  
Db 3063 ATCAGATGATGTGCGGCCGACCTGGAGTCTGCACGAGCACCTGGGTGCTCGTTGGCGG 3122  
Qy 641 ValLeuAlaAlaLeuAlaAlaTyrCysLeuSerThrGlyCysValValIleValGlyArg 660  
Db 3123 GTCCTGGCTGCTTTGGCCCGATTCCTCTCAACAGGCTGCTGTGATGTCAGTGGCAGG 3182  
Qy 661 IleValLeuSerGlyLysProAlaIleIleProAspArgGluValLeuTyrArgGluPhe 680  
Db 3183 GTCGCTCTTCTCGGGAAGCCGCAATCATACCTGACAGGAACTCTCTTACCGAGATTTC 3242  
Qy 681 AspGluMetGluGluCys 686  
Db 3243 GATGAGATGGAAGAGTGC 3260

## RESULT 9

US-08-444-818-88  
; Sequence 88, Application US/08444818  
; Patent No. 6150087  
; GENERAL INFORMATION:  
; APPLICANT: Chien, David Y.  
; APPLICANT: Rutter, William J.  
; TITLE OF INVENTION: NANBV Diagnostics and Vaccines  
; NUMBER OF SEQUENCES: 777  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Chiron Corporation  
; STREET: 4560 Horton Street  
; CITY: Emeryville  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 94608-2916  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/444,818  
; FILING DATE:  
; CLASSIFICATION: 424  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/08/403,590  
; FILING DATE: 14-MAR-1995  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Harbin, Alisa A.  
; REGISTRATION NUMBER: 33,895  
; REFERENCE/DOCKET NUMBER: 0110.002  
; TELEPHONE: (508)359-3876  
; TELEFAX: (508)359-3885  
; INFORMATION FOR SEQ ID NO: 88:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 8316 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: cDNA  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: 1..8316  
; US-08-444-818-88  
Alignment Scores:  
Pred. No.: 0 Length: 8316  
Score: 3570.00 Matches: 671  
Percent Similarity: 99.27% Conservative: 10  
Best Local Similarity: 97.81% Mismatches: 5



Query Match:	98.67%	Indels:	0
DB:	3	Gaps:	0
US-09-930-591-2 (1-686) x US-08-444-818-88 (1-8316)			
Qy	1	MetAlaProIleThrAlaTyrAlaGlnGlnThrArgGlyLeuLeuGlyCysIleIleThr	20
Db	2734	CTGGCGCCCATCACGGGTACGCCACAGACAGAGGGCCCTCTTAGGGTGCATAATCACC	2793
Qy	21	SerLeuThrGlyArgAspLysAsnGlnValcluglyGluValGlnIleValSerThrAla	40
Db	2794	AGCCTAACTGGCGGGACAAAACCAAGTGGAGGTGAGGTCCAGATTGTGTCAACTGCT	2853
Qy	41	AlaGlnThrPheLeuAlaThrCysIleAsnGlyValCysTrpThrValTyrHisGlyAla	60
Db	2854	GCCCAAACTTCTCGCAACGTGCATCAATGGGGTGTGCTGGACTGTCTACCAACGGGGCC	2913
Qy	61	GlyThrArgThrIleAlaSerProLysGlyProValIleGlnMetTyrThrAsnValAsp	80
Db	2914	GGACAGGAGGACCATCGCGTCCACCAAGGTCTGTCTATCCAGATGTATACCAATGTAGAC	2973
Qy	81	GlnAspLeuValGlyTrpProAlaProGlnGlyAlaArgSerLeuThrProCysThrCys	100
Db	2974	CAAGACCTTGTGGCTGCCCGCTCCGCAAGTAGCCGCTCATTTGACACACCTGCACCTGC	3033
Qy	101	GlySerSerAspLeuTyrLeuValThrArgHisAlaAspValIleProValArgArgArg	120
Db	3034	GGCTCTCTCGACCTTTACTTGTGTCCAGAGCACGCCGATGTTCATTCCTCGTCGCCCGCGG	3093
Qy	121	GlyAspGlyArgGlySerLeuLeuSerProArgProIleSerTyrLeuLysGlySerSer	140
Db	3094	GGTGATAGCAGGGGACGCTGTGTGCCCGCGGCCCATTTCTACTTGAAGAGCTCTCTCG	3153
Qy	141	GlyGlyProLeuLeuCysProAlaGlyHisAlaValGlyIlePheArgAlaAlaValCys	160
Db	3154	GGGGGTCCGCTGTGTGCCCGCGGGGACGCGCTGGGCATATTTAGGGCGCGCGTGTGC	3213
Qy	161	ThrArgGlyValAlaLysAlaValAspPheIleProValGluSerLeuGluThrThrMet	180
Db	3214	ACCCGTGGAGTGGCTAAGCGCGGTGGACTTTATTCCTGTGGAGAACCTTAGAGACAAACCATG	3273
Qy	181	ArgSerProValPheSerAspAsnSerSerProProAlaValProGlnSerTyrGlnVal	200
Db	3274	AGGTCCCGGTGTTCACGGATAACTCTCTCCACAGTAGTGTGCCCCAGAGCTTCACAGTG	3333
Qy	201	AlaHisLeuHisAlaProThrGlySerGlyLysSerThrLysValProAlaAlaTyrAla	220
Db	3334	GCTCACCTCCATGTCTCCACAGGCAGCGGCAAAAGCACCAAGGTCCCGGCTGCATATGCA	3393
Qy	221	AlaGlnGlyTyrLysValLeuValIleAsnProSerValAlaAlaThrMetGlyPheGly	240
Db	3394	GCTCAGGGCTATAAGGTGCTAGTACTCAACCCCTCTGTGTCTGCAACACTGGGGCTTTGGT	3453
Qy	241	AlaTyrMetSerLysAlaHisGlyIleAspProAsnIleArgThrGlyValArgThrIle	260
Db	3454	GCTTACATGTCCAAGGCTCATGGATCGATCCTAACATCAGGACCGGGGTGAGAACAAATT	3513
Qy	261	ThrThrGlySerProIleThrTyrSerThrTyrGlyLysPheLeuAlaAspGlyGlyCys	280
Db	3514	ACCACTGGCAGCCCCATCAGTACTCCACCTTACGGCAAGTTCTCTTGTGCCGACGGCGGTGC	3573
Qy	281	SerGlyGlyAlaTyrAspIleIleIleCysAspGluCysHisSerThrAspAlaThrSer	300
Db	3574	TCGGGGGGCGGTTATGACATAAATTTGTGACGAGTGGCACTCCACGGATGCCACATCC	3633
Qy	301	IleLeuGlyIleGlyThrValLeuAspGlnAlaGluThrAlaGlyAlaArgLeuThrVal	320
Db	3634	ATCTTGGGCATCGCACTGTCTTGCACCAAGCAGAGACTCGGGGGCGAGACTGTTGTG	3693
Qy	321	LeuAlaThrAlaThrProProGlySerValThrValProHisProAsnIleGluGluVal	340
Db	3694	CTCGCACCGCACCCCTCGGGCTCGGTCTACTGTGCGCCATCCATCCACATCGAGAGGTT	3753

Qy	341	AlaLeuSerThrThrGlyGluIleProPheTyrGlyLeuAlaIleProLeuGluAlaIle	360
Db	3754	GCTCTGTCCACCGAGAGATCCCTTTTACGGAAGGCTATCCCCCTCGAAGTAATC	3813
Qy	361	LysGlyGlyVArgHisLeuIlePheCysHisSerLysLysLysCysAspGluLeuAlaAla	380
Db	3814	AAGGGGGGAGACATCTCATCTTCTGTCTCATTAAGAAGAAGTGCAGCACTCGCGCA	3873
Qy	381	LysLeuValAlaLeuGlyValAsnAlaValAlaTyrTyrArgGlyLeuAspValSerVal	400
Db	3874	AAGCTGCTGCATTTGGGCATCAATGCGTGGCCTACTACCGCGTCTTGACGTGTCGTC	3933
Qy	401	IleProThrSerGlyAspValValValAlaAlaThrAspAlaLeuMetThrGlyPheThr	420
Db	3934	ATCCGACACAGCGCGATGTTGTCTGTGTGGCAACCGATGCCCTCATGACCGGCTATACC	3993
Qy	421	GlyAspPheAspSerValIleAspCysAsnThrCysValThrGlnThrValAspPheSer	440
Db	3994	GGCGACTTCGACTCGGTGTAGTAGCTGCAATACGTGTGTACCCAGACAGTCGATTTTCAGC	4053
Qy	441	LeuAspProThrPheThrIleGluThrIleThrLeuProGlnAspAlaValSerArgThr	460
Db	4054	CTTGACCTTACCTTCACATTTGAGACAAATCACGCTCCCCAGAGATGCTGCTCCCGCACT	4113
Qy	461	GlnArgGlyArgThrGlyArgGlyLysProGlyIleTyrArgPheValAlaProGly	480
Db	4114	CAACGTCCGGGCGAGACTGGCAGGGGGAAGCCAGGCACTCTACAGATTTGTGGCACCGGG	4173
Qy	481	GluArgProSerGlyMetPheAspSerSerValLeuCysGluCysTyrAspAlaGlyCys	500
Db	4174	GAGCGCCCTCCCGCATGTTTCGACTCGTCCGTCTCTGTGAGTGCTATGACGAGGCTGT	4233
Qy	501	AlaTrpTyrGluLeuThrProAlaGluThrThrValArgLeuArgAlaTyrMetAsnThr	520
Db	4234	GCTTGTATGAGCTCACGCCCGCCGAGACTACAGTTAGGCTACGACGTCATGAACACC	4293
Qy	521	ProGlyLeuProValCysGlnAspHisLeuGluPheTrpGluGlyValPheThrGlyLeu	540
Db	4294	CCGGGGCTTCCCGTGTGCCAGGACCATCTTTGAATTTTGGAGGGCGTCTTTACAGGCCTC	4353
Qy	541	ThrHisIleAspAlaHisPheLeuSerGlnThrLysGlnSerGlyGluAsnLeuProTyr	560
Db	4354	ACTCATATAGATGCCACTTCTATCCGACAAAGCAGAGTGGGAGAACCTTCCTCTTAC	4413
Qy	561	LeuValAlaTyrGlnAlaThrValCysAlaAtrqAlaGlnAlaProProSerTrpAsp	580
Db	4414	CTGTGACGTACCAAGCCACCGTGTGCGTGTAGGGCTCAAGCCCTCCCCCATCGTGGGAC	4473
Qy	581	GlnMetTrpLysCysLeuIleArgLeuLysProThrLeuHisGlyProThrProLeuLeu	600
Db	4474	CAGATGTGGAGTGTGATTCCGCTCAAGCCACACCTCCATGGGCCAACACCCCTGCTA	4533
Qy	601	TyrArgLeuGlyAlaValGlnAsnGluValThrLeuThrHisProValThrLysTyrIle	620
Db	4534	TACAGACTGGCGCTGTTTCAGATGAATCACCCTGACGCACCCTGACCAAAATACATC	4593
Qy	621	MetThrCysMetSerAlaAspLeuGluValValThrSerThrTrpValLeuValGlyGly	640
Db	4594	ATGACATGCATGTGGCCGACCTGGAGGTGCTCACGAGCACCTGGGTGCTGCTGTGGCGGC	4653
Qy	641	ValLeuAlaAlaLeuAlaAlaTyrCysLeuSerThrGlyCysValValIleValGlyArg	660
Db	4654	GTCTGGCTGCTTTGGCCGGGTATTGCCTGTCAACAGGCTGCTGTGTCATAGTGGGCGG	4713
Qy	661	IleValLeuSerGlyLysProAlaIleIleProAspArgGluValLeuTyrArgGluPhe	680
Db	4714	GTCTGCTTGTCCGGGAAGCCGCAATCATACCTTGCAGGGAAGTCTCTTACGAGAGTTTC	4773
Qy	681	AspGluMetGluGluCys	686
Db	4774	GATGAGATGGAAGATGC	4791

```

US-08-444-818-122
; Sequence 122, Application US/08444818
; Patent No. 6150087
; GENERAL INFORMATION:
; APPLICANT: Chien, David Y.
; APPLICANT: Rutter, William J.
; TITLE OF INVENTION: NANBV Diagnostics and Vaccines
; NUMBER OF SEQUENCES: 777
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Chiron Corporation
; STREET: 4560 Horton Street
; CITY: Emeryville
; STATE: CA
; COUNTRY: USA
; ZIP: 94608-2916
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/444,818
; FILING DATE:
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/403,590
; FILING DATE: 14-MAR-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Harbin, Alisa A.
; REGISTRATION NUMBER: 33,895
; REFERENCE/DOCKET NUMBER: 0110.002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (508)359-3876
; TELEFAX: (508)359-3885
; INFORMATION FOR SEQ ID NO: 122:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 9185 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cdna
; US-08-444-818-122

```

[illegible]

US-09-930-591-2 (1-686) x US-08-444-818-122 (1-9185)

Qy		1	MetaAlaProfileThrAlaTyrAlaGlnGlnThrArgGlyLeuLeuGlyCysIleIleThr	20
	:	:	:	:
Db		3395	CTGGCGCCCATACGGCGTAGCCAGACAGAGGGCGCTCTCTAGGTGCATAATCAC	3454
Qy		21	SerLeuThrGlyArgAspLysAsnGlnValGlucGlyGluValGlnIleValSerThrAla	40
	:	:	:	:
Db		3455	AGCCTAACTGGCGGGHCAAAAACAAGTGAGGGTGAGGTCCNGATTGTGTCAATGCT	3514
Qy		41	AlaGlnThrPheLeuAlaThrCysIleAsnGlnValCysTrpThrValTyrHisGlyAla	60
	:	:	:	:
Db		3515	GCCCAAACCTTCCTGGCAAGGTGCATCAATGGGGTGCTCGAGTGTCTACCACGGGGCC	3574
Qy		61	GlyThrArgThrIleAlaSerProLysGlyProValIleGlnMetTyrThrAsnValasp	80
	:	:	:	:
Db		3575	GGAAAGAGGACCATCGCGTACCACAGGGTCCTGTCAICAGATGATATCAATGTAAC	3634
Qy		81	GlnAspLeuValGlyTyrProAlaProGlnGlnYalaAargSerLeuThrProCysThrCys	100
	:	:	:	:
Db		3635	CAGAAGCTTTGGGCTGGCCCGGTCCGAAGGTAGCGCGTCATTGACACCGCTGCATCTGC	3694

Qy	101	GlySerSerAspLeuTyrLeuValThrArgHisAlaAspValIleProValArgArg	120
Db	3695	GGTCTCTCGACCTTACTCTGGTGCAGAGGACGCGCATGTCAATCCCGTGC	3754
Qy	121	GlyAspGlyArgGlySerLeuLeuSerProArgProIleSerTyrLeuLysGlySerSer	140
Db	3755	GGTGATAGCAGGGGACGCTGCTGTGCGCCGCGCCCATTTCTTCTTGAAGGCTCTCTCG	3814
Qy	141	GlyGlyProLeuLeuCysProAlaGlyHisAlaValGlyIlePheArgAlaAlaValCys	160
Db	3815	GGGGTCCGTGTGTGTCGCCCGGGGACGCGTGGGCATATTAGGCGCGGTGTGC	3874
Qy	161	ThrArgGlyValAlaLysAlaValAspPheIleProValGluLuserLeuGluThrThrMet	180
Db	3875	ACCCGTGGAGTGGCTAAGGCGGTGGACTTTATCCCTGTGGAGAACCTAGAGACAACCATG	3934
Qy	181	ArgSerProValPheSerAspAsnSerSerProAlaValProGlnSerTyrGlnVal	200
Db	3935	AGGTCCCGGTGTTCAGGATAACTCTCTCCACCAAGTAGTGCCCCAGAGCTTCCAGGTG	3994
Qy	201	AlaHisLeuHisAlaProThrGlySerGlyLysSerThrLysValProAlaIatYrAla	220
Db	3995	GCTCACCTCATGCTCCACAGCAGCGGCAAAAGCACCAAGGTCCCGGTGCATATGCA	4054
Qy	221	AlaGlnGlyTyrLysValLeuValLeuAsnProSerValAlaAlaThrMetGlyPheGly	240
Db	4055	GCTCAGGGCTATAAGGTGCTAGTACTCAACCCCTCTGTGCTGCAACACTGGGCTTTGGT	4114
Qy	241	AlaTyrMetSerLysAlaHisGlyIleAspProAsnIleArgThrGlyValArgThrIle	260
Db	4115	GCTTATCATGTCCAAGGCTCATGGGATCGATCTCTTAACATCAGGACCGGGGTGAGAACAAAT	4174
Qy	261	ThrThrGlySerProIleThrTyrSerThrTyrGlyLysPheLeuAlaAspGlyLysCys	280
Db	4175	ACCACTGGCAGCCCCATCAGTACTCCACTAGGCGAAGTCTCTGCCGACGCGGGGTGC	4234
Qy	281	SerGlyGlyAlaTyrAspIleIleLysAspGluCysHisSerThrAspAlaThrSer	300
Db	4235	TCGGGGCGGGTTATGACATAATAATTTGTGACAGTGCACCTCCACGGATGCCACATCC	4294
Qy	301	IleLeuGlyIleGlyThrValLeuAspGlnAlaGluThrAlaGlyAlaArgLeuThrVal	320
Db	4295	ATCTTGGGCATCGGCATGTCTTGTGACCAAGCAGAGACTCGCGGGCGAGACTGGTTGTG	4354
Qy	321	LeuAlaThrAlaThrProGlySerValThrValProHisProAsnIleGluGluVal	340
Db	4355	CTGCCACCGCCACCCCTCGGGCTCGTCACTGTGCCCCCATCCCAACATCGAGGAGGTT	4414
Qy	341	AlaLeuSerThrThrGlyGluIleProPheTyrGlyLysAlaIleProLeuGluAlaIle	360
Db	4415	GCTCTGTCCACCCCGGAGAGATCCCTTTTACGGCAAGGCTATCCCTCTCGAAGTAATC	4474
Qy	361	LysGlyGlyArgHisIleIlePheCysHisSerLysLysLysCysAspGluLeuAlaAla	380
Db	4475	AAGGGGGGAGACATCTCATCTTCTGTCAATCAAAGAAGAAGTGCAGCAACTCGCGCGCA	4534
Qy	381	LysLeuValAlaLeuGlyValAsnAlaValAlaIatTyrArgGlyLeuAspValSerVal	400
Db	4535	AAGTGTGTGCATTTGGGCATCAATGCGGGGCTCTACTCCGGGCTTTGACGTGTCGCTC	4594
Qy	401	IleProThrSerGlyAspValValValAlaIatThrAspAlaLeuMetThrGlyPheThr	420
Db	4595	ATCCCGACCGCGCGATGTTGTGTCGTGGCAACCGATGCCCTCATGACCGGCTATACC	4654
Qy	421	GlyAspPheAspSerValIleAspCysAsnThrCysValThrGlnThrValAspPheSer	440
Db	4655	GGCGACTTCTGACTCGGTGTAGACTCGCAATACGTGTGTCAACCCAGACAGTTCGATTTT	4714
Qy	441	LeuAspProThrPheThrIleGluThrIleThrLeuProGlnAspAlaValSerArgThr	460
Db	4715	CTTGACCCCTACCTTTCACCATTTGAGACAAATCACGCTCCCCAGAGTGTGTCTCCCGACT	4774
Qy	461	GlnArgArgGlyArgThrGlyArgGlyLysProGlyLysIleTyrArgPheValAlaProGly	480

Db 4775 CAACGTGGGCGAGACTGGCAGGGGAGCCAGGCATCTACAGATTGTGGCCACCGGG 4834  
Qy 481 GluArgProSerGlyMetPheAspSerValLeuGluCysTyrAspAlaGlyCys 500  
Db 4835 GAGCGCCCTCCGCGATGTCGATCGCTCCCTCTGTGAGTCTATGACGCGAGGTGT 4894  
Qy 501 AlaTrpTyrGluLeuThrProAlaGluThrThrValArgLeuArgAlaTyrMetAsnThr 520  
Db 4895 GCTTGGTATAGCTCACGCGCGCGAGACTACAGTTAGGCTACGAGCGTACATGAACACC 4954  
Qy 521 ProGlyLeuProValCysGlnAspHisLeuGluPheTrpGluGlyValPheThrGlyLeu 540  
Db 4955 CCGGGGCTTCCTGGTGGCAGGACCATCTTGAATTTTGGGAGGGCGCTTTTACAGGGCTC 5014  
Qy 541 ThrHisLeuAspAlaHisPheLeuSerGlnThrLysGlnSerGlyGluAsnLeuProTyr 560  
Db 5015 ACTCATATAGATGCCACCTTCTATCCAGACAAGCAGAGTGGGGAGAACCTTCTCTTAC 5074  
Qy 561 LeuValAlaTyrGlnAlaThrValCysAlaArgAlaGlnAlaProProSerTrpAsp 580  
Db 5075 CTGGTAGCGTACCAAGCCACCGTGGCGTAGGGCTCAAGGCCCTCCCCCATCGTGGGAC 5134  
Qy 581 GlnMetTrpLysCysLeuLeuArgLeuLysProThrLeuHisGlyProThrProLeuLeu 600  
Db 5135 CAGATGTGGAAGTGTGTTGATTGGCTCAAGGCCACCTCCATGGGCCCAACACCCCTGCTA 5194  
Qy 601 TyrArgLeuGlyAlaValGlnAsnGluValThrLeuThrHisProValThrLysTyrIle 620  
Db 5195 TACAGACTGGCGCGTGTTCAGAAATGAAATACCTGAGCGACCCAGTCACCAATATATC 5254  
Qy 621 MetThrCysMetSerAlaAspLeuGluValValThrSerThrTrpValLeuValGlyGly 640  
Db 5255 ATGACATGCATGTCCGGCGACCTGGAGTGTCTACGAGCACCTGGGTGCTGTTGGCGGC 5314  
Qy 641 ValLeuAlaAlaLeuAlaAlaTyrCysLeuSerThrGlyCysValValIleValGlyArg 660  
Db 5315 GTCTGTGCTGTTTGGCGCGTATTTGCTGTCAACAGGCTGCGTGTGCATAGTGGCAGG 5374  
Qy 661 IleValLeuSerGlyLysProAlaIleProAspArgGluValLeuTyrArgGluPhe 680  
Db 5375 GTCTGTGTTCCGGAGCGCGCAATCATCTCAGAGGAAAGTCTCTACCGAGAGTTC 5434  
Qy 681 AspGluMetGluGluCys 686  
Db 5435 GATCAGATGGAAGAGTGC 5452

RESULT 11

US-08-444-818-123/c  
; Sequence 123, Application US/08444818  
; Patent No. 6150087  
; GENERAL INFORMATION:  
; APPLICANT: Chien, David Y.  
; APPLICANT: Ruteer, William J.  
; TITLE OF INVENTION: NANB Diagnostics and Vaccines  
; NUMBER OF SEQUENCES: 777  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Chiron Corporation  
; STREET: 4560 Horton Street  
; CITY: Emeryville  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 94608-2916  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/444,818  
; FILING DATE:  
; CLASSIFICATION: 424  
; PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/08/403,590  
FILING DATE: 14-MAR-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Harbin, Alisa A.  
REGISTRATION NUMBER: 33,895  
REFERENCE/DOCKET NUMBER: 0110.002  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (508)359-3876  
TELEFAX: (508)359-3885  
INFORMATION FOR SEQ ID NO: 123:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 9185 base pairs  
TYPE: nucleic acid  
STRADEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cdna  
ANTI-SENSE: YES  
US-08-444-818-123  
Alignment Scores:  
Pred. No.: 0 Length: 9185  
Score: 3570.00 Matches: 671  
Percent Similarity: 99.27% Conservative: 10  
Best Local Similarity: 97.81% Mismatches: 5  
Query Match: 98.67% Indels: 0  
DB: 3 Gaps: 0  
US-09-930-591-2 (1-686) x US-08-444-818-123 (1-9185)  
Qy 1 MetAlaProIleThrAlaTyrAlaGlnGlnThrArgGlyLeuLeuGlyCysIleIleThr 20  
Db 5791 CTGCGGCCATCATCGCGGTACGCCACGAGCAAGGGCCCTCCTAGGGTGCATATCACC 5732  
Qy 21 SerLeuThrGlyArgAspLysAsnGlnValGluGlyGluValGlnIleValSerThrAla 40  
Db 5731 AGCTTAACCTGGCCGGGCAAAAACCAAGTGGAGGTGAGTCCAGATTGTGTCAACTGCT 5672  
Qy 41 AlaGlnThrPheLeuAlaThrCysIleAsnGlyValCysTrpThrValTyrHisGlyVala 60  
Db 5671 GCCCAAAACCTTCTGGCAACGTGATCAATGGGGTGTGCTGGACTGTCTACACCGGGGCC 5612  
Qy 61 GlyThrArgThrIleAlaSerProLysGlyProValIleGlnMetTyrThrAsnValAsp 80  
Db 5611 GGAACGAGGACCATCGCGTCACCAAGGGTCTGTCTATCCAGATGTATACCAATGTAGAC 5552  
Qy 81 GlnAspLeuValGlyTrpProAlaProGlnGlnValaArgSerLeuThrProCysThrCys 100  
Db 5551 CAAGACCTTGTGGGTGGCCCGCTCCGCAAGGTAGCCGCTCATTTGACACCCCTGCACCTGC 5492  
Qy 101 GlySerSerAspLeuTyrLeuValThrArgHisAlaAspValIleProValArgArg 120  
Db 5491 GGCTCCTCGGACCTTTACCTGGTTCAGAGGACGCGCATGTCTATCCGTCGCGCGGG 5432  
Qy 121 GlyAspGlyArgGlySerLeuLeuSerProAspProIleSerTyrLeuLysGlySerSer 140  
Db 5431 GGTATAGACAGGGGCGACCTGTGTGCGCCGCGCCATTTCTACTTCTAAAGGCTCCTCG 5372  
Qy 141 GlyGlyProLeuLeuCysProAlaGlyHisAlaValGlyIlePheArgAlaValCys 160  
Db 5371 GGGGGTCCGCTGTGTGTCGCCCGCGGGGACGCGCGTGGGCATATTTAGGGCGCGGTGTGC 5312  
Qy 161 ThrArgGlyValAlaLysAlaValAspPheIleProValGluSerLeuGluThrThrMet 180  
Db 5311 ACCGTGGAGTGGCTAAGCGGTGGACTTTATCCCTGTGGAGAACCTTAGACACCATG 5252  
Qy 181 ArgSerProValPheSerAspAsnSerSerProProAlaValProGlnSerTyrGlnVal 200  
Db 5251 AGGTCCCGGTGTTCACGGATAACTCTCTCCACAGTAGTGTGCCACAGAGCTTCCAGGTG 5192  
Qy 201 AlaHisLeuHisAlaProThrGlySerGlyLysSerThrLysValProAlaAlaTyrAla 220  
Db 5191 GCTCACCTCCATGCTCCACAGGACGCGGCAAGGACCAAGGTCGCCGTGCATATGCA 5132

```
QY 221 AlaGlnGlyTyrLysValLeuValLeuAsnProSerValAlaAlaThrMetGlyPheGly 240
Db 5131 GCTCAGGCTATAAGGTGTAGTACTCAACCCCTCTGTTGCTGCAACACTGGGCTTTGCT 5072
QY 241 AlaTyrMetSerLysAlaHisGlyIleAspProAsnIleArgThrGlyValArgThrIle 260
Db 5071 GCTTACATGTCCAAGGCTCATGGATCGATCTTAACATCAGGACCGGGGTGAACAATT 5012
QY 261 ThrThrGlySerProIleThrTyrSerThrTyrGlyLysPheLeuAlaAspGlyGlyCys 280
Db 5011 ACCACTGGACGCCCATCAGTACTCCACTACGGCAAGTTCCTTCCGACGCGGGTGC 4952
QY 281 SerGlyGlyAlaTyrAspIleIleCysAspGluCysHisSerThrAspAlaThrSer 300
Db 4951 TCGGGGGCGGTTATGACATAAATTGTGACGAGTGCACATCCACGGATGCCACATCC 4892
QY 301 IleLeuGlyIleGlyThrValLeuAspGlnAlaGluThrAlaGlyAlaArgLeuThrVal 320
Db 4891 ATCTTGGGCAATCGGCATCTGCTTGAACCAAGCAGAGACTCGGGGGCGAGACTGGTGTG 4832
QY 321 LeuAlaThrAlaThrProGlySerValThrValProHisProAsnIleGluGluVal 340
Db 4831 CTGCGCACCGCCACCCCTCGGGCTCGTCTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4772
QY 341 AlaLeuSerThrThrGlyGluIleProPheTyrGlyLysAlaIleProLeuGluAlaIle 360
Db 4771 GCTCTGTCCACCCGAGAGATCCCTTTTACGGCAAGCTATCCCTCCGAAAGTAATC 4712
QY 361 LysGlyGlyArgHisLeuIlePheCysHisSerLysLysLysCysAspGluLeuAlaAla 380
Db 4711 AAGGGGGGAGACATCTCATCTTCTGTCAATCAAGAAGAAGTGCAGCAACTCGCGCA 4652
QY 381 LysLeuValAlaLeuGlyValAsnAlaValAlaTyrArgGlyLeuAspValSerVal 400
Db 4651 AAGTGTGTCGATTTGGGCATCAATCGCGTGGCTACTACCGCGTCTTGACGTCCGTC 4592
QY 401 IleProThrSerGlyAspValValAlaThrAspAlaLeuMetThrGlyPheThr 420
Db 4591 ATCCGACACCGCGGATTTGCTGCTGCGCAACCGATGCCCTCATGACCGGCTATACC 4532
QY 421 GlyAspPheAspSerValIleAspCysAsnThrCysValThrGlnThrValAspPheSer 440
Db 4531 GCGGACTTCGACTCGGTGATAGACTGCAATACGTGTGTACCCAGACAGCTCGATTTTCAGC 4472
QY 441 LeuAspProThrPheThrIleGluThrIleThrLeuProGlnAspAlaValSerArgThr 460
Db 4471 CTGACCTACCTTTCACATTTAGACAAATCAGCTCCCCAGATGCTGTCTCCGCACT 4412
QY 461 GlnArgArgGlyArgThrGlyArgGlyLysProGlyIleTyrArgPheValAlaProGly 480
Db 4411 CAACGTCGGGCGAGGACTGGCAGGGGAGCCAGGCATCTACAGATTTGTGGCACCGGG 4352
QY 481 GluArgProSerGlyMetPheAspSerSerValLeuCysGluCysTyrAspAlaGlyCys 500
Db 4351 GAGCGCCCTCCGCGATGTTTCGACTCGCTCGCTCTCTGTGAGTGTATGACGAGGCTGT 4292
QY 501 AlaTyrThrGluLeuThrProAlaGluThrThrValArgLeuArgAlaTyrMetAsnThr 520
Db 4291 GCTTGGTATGAGCTCACGCCCGCGAGACTACAGTTAGGCTACGAGCTACATGAACACC 4232
QY 521 ProGlyLeuProValCysGlnAspHisLeuGluPheThrPheGluGlyValPheThrGlyLeu 540
Db 4231 CCGGGGCTTCCGCTGCCAGGACCATCTTGAATTTTGGAGGGCGTCTTTTACAGGCTC 4172
QY 541 ThrHisIleAspAlaHisPheLeuSerGlnThrLysGlnSerGlyGluAsnLeuProTyr 560
Db 4171 ACTCATATAGTGCCTTCTATCCAGACAAAGCAGAGTGGGGAGAACCTTCTCTTAC 4112
QY 561 LeuValAlaTyrGlnAlaThrValCysAlaArgAlaGlnAlaProProSerThrAsp 580
Db 4111 CTGGTAGGTACCAAGCAACCGTGTGCTAGGCTCAGGCTCAGGCCCTCCCCCATCGTGGAC 4052
QY 581 GlnMetTrpLysCysLeuIleArgLeuLysProThrLeuHisGlyProThrProLeuLeu 600
```

```
Db 4051 CAGATGTGGAAGTGTGTTGATTGCGCTCAAGGCCACCCCTCCATGGGCCAACCCCTGTCTA 3992
QY 601 TyrArgLeuGlyAlaValAlaGlnAsnGluValThrLeuThrHisProValThrLysTyrIle 620
Db 3991 TACAGACTGGGCGCTGTTTCAGAAATCAATCACCCTGAGCGACCCAGTACCAAAATACATC 3932
QY 621 MetThrCysMetSerAlaAspLeuGluValValThrSerThrTrpValLeuValGlyGly 640
Db 3931 ATGACATCATGCTCGCCGACCTGGAGTCTGTCACGAGCACCTGGGTGCTGTTGGCGGC 3872
QY 641 ValLeuAlaLeuAlaAlaTyrCysLeuSerThrGlyCysValValIleValGlyArg 660
Db 3871 GTCTGTGCTGCTTTCGCGCGTATTTCCTGTCAACAGGCTGCGTGTCTATAGTGGCAGG 3812
QY 661 IleValLeuSerGlyLysProAlaIleIleProAspArgGluValLeuTyrArgGluPhe 680
Db 3811 GTCTGTCTGTCGGGAAGCGGCAATCATACCTGACAGGAAGTCTCTTACCGAGATTC 3752
QY 681 AspGluMetGluGluCys 686
Db 3751 GATGAGATGAAGAGTGC 3734
RESULT 12
US-09-881-239-2
; Sequence 2, Application US/09881239
; Patent No. 6630298
; GENERAL INFORMATION:
; APPLICANT: CHIEN, David Y.
; APPLICANT: ARCANDEL, Phillip
; APPLICANT: TANDESKIE, Laura
; APPLICANT: GEORGE-NASCIMENTO, Carlos
; APPLICANT: COIT, Doris
; APPLICANT: MEDINA-SELBY, Angelica
; TITLE OF INVENTION: HCV ANTIGEN/ANTIBODY COMBINATION ASSAY
; FILE REFERENCE: 2302-16073 / PP16073.003
; CURRENT APPLICATION NUMBER: US/09/881,239
; CURRENT FILING DATE: 2001-06-14
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 2058
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:
; OTHER INFORMATION: representative NS3/4a conformational antigen
; NAME/KEY: CDS
; LOCATION: (1)..(2058)
US-09-881-239-2
Alignment Scores:
Pred. No.: 0 Length: 2058
Score: 3565.00 Matches: 671
Percent Similarity: 99.13% Conservatives: 9
Best Local Similarity: 97.81% Mismatches: 6
Query Match: 98.54% Indels: 0
DB: 4 Gaps: 0
US-09-930-591-2 (1-686) x US-09-881-239-2 (1-2058)
QY 1 MetAlaProIleThrAlaTyrAlaGlnGlnThrArgGlyLeuLeuGlyCysIleIleThr 20
Db 1 ATGGCGCCCATCACGGCGTACGCCACAGACAGCAAGGGGCTCTCCATGGGTGCAATAATCACC 60
QY 21 SerLeuThrGlyArgAspLysGlnValGluGlyGluValGlnIleValSerThrAla 40
Db 61 AGCTTAACCTGGCGGGACAAAACCAAGTGGAGGGTGAAGTCCAGATTGTGTCACTGCT 120
QY 41 AlaGlnThrPheLeuAlaThrCysIleAsnGlyValCysTrpThrValTyrHisGlyAla 60
Db 121 GCCCAACCTTCTCTGCAACGTCATCAATGGGGTGTGTGGACTGTCTTACCGAGGGCC 180
```

Qy 61 GlyThrArgThrIleAlaSerProLysGlyProValIleGlnMetTyrThrAsnValAsp 80  
Db 181 GGAACGAGGACCATCGCGTCAACCAAGGGTCTGTGCATCCAGATGATACCAATGTAGAC 240  
Qy 81 GlnAspLeuValGlyTyrProAlaProGlnGlyAlaArgSerLeuThrProCysThrCys 100  
Db 241 CAAGACCTTGTGGCTGGCGTCCGCAAGGTAGCCGATCATTTGACACCTCGCATTCG 300  
Qy 101 GlySerSerAspLeuTyrLeuValThrArgHisAlaAspValIleProValArgArgArg 120  
Db 301 GGCCTCTCGACCTTTACCTGGTCAGGAGGACGCGGATGTCATTCCTGGCGCGGGG 360  
Qy 121 GlyAspGlyArgGlySerLeuLeuSerProArgProIleSerTyrLeuLysGlySerSer 140  
Db 361 GGTGATAGCAGGGGAGCGCTGTGTGCGCCCGGCCCATTTCTACTTGAAGGCTCTCTCG 420  
Qy 141 GlyGlyProLeuLeuCysProAlaGlyHisAlaValGlyIlePheArgAlaAlaValCys 160  
Db 421 GGGGTCCTGCTGTGTGCGCGGGGACGCGGTGGGCATATTTAGGGCGCGGTGTGC 480  
Qy 161 ThrArgGlyValAlaLysAlaValAspPheIleProValGluSerLeuGluThrThrMet 180  
Db 481 ACCGTGGAGTGGCTAAGCGGTGGACTTTATCTCTGTGGAGAACCTAGAGACCAACATG 540  
Qy 181 ArgSerProValPheSerAspAsnSerSerProProAlaValProGlnSerTyrGlnVal 200  
Db 541 AGGTCCCGGGTGTTCACGGATAACTCTCTCCACCCAGTGTGCCCCAGAGCTTCAGGTG 600  
Qy 201 AlaHisLeuHisAlaProThrGlySerGlyLysSerThrLysValProAlaAlaTyrAla 220  
Db 601 GCTCACTCTCCATGCTCCACAGGAGCGGCAAAAGCACCAAGGTCCCGGTGTCATATGCA 660  
Qy 221 AlaGlnGlyTyrLysValLeuValLeuAsnProSerValAlaAlaThrMetGlyPheGly 240  
Db 661 GCTCAGGGCTATAGGTCTAGTACTCAACCCCTCTGTGTGTGCAACTCTGGGCTTTGGT 720  
Qy 241 AlaTyrMetSerLysAlaHisGlyIleAspProAsnIleArgThrGlyValArgThrIle 260  
Db 721 GCTTACATGTCCAAGGCTCATGGATCGATCTTAACATCAGGACCGGGGTGAGAACAT 780  
Qy 261 ThrThrGlySerProIleThrTyrSerThrTyrGlyLysPheLeuAlaAspGlyLysCys 280  
Db 781 ACCACTGGCAGCCCATCAGCTACTCCACCTACCGCAAGTTCTTGGCGAGCGGGGTGC 840  
Qy 281 SerGlyGlyAlaTyrAspIleIleCysAspGluCysHisSerThrAspAlaThrSer 300  
Db 841 TCGGGGGCGGCTTATGACATATATTTGTGACAGGTGCCACTCCAGATGCCACATCC 900  
Qy 301 IleLeuGlyIleGlyThrValLeuAspGlnAlaGluThrAlaGlyAlaArgLeuThrVal 320  
Db 901 ATCTTGGGCAATTGGCACTGTCTTGACCAAGCAGAGACTGCGGGGGCGAGACTGGTTGTG 960  
Qy 321 LeuAlaThrAlaThrProProGlySerValThrValProHisProAsnIleGluGluVal 340  
Db 961 CTCGCCACCGCACCCCTCCGGGTCTCGTCACTGTGCCCATCCCAACATCGAGGAGTT 1020  
Qy 341 AlaLeuSerThrThrGlyGluIleProPheTyrGlyLysAlaIleProLeuGluAlaIle 360  
Db 1021 GCTCTGCCACCGAGAGATCCCTTTTACGGCAAGGCTATCCCCCTCGAAGTAACT 1080  
Qy 361 LysGlyGlyArgHisLeuIlePheCysHisSerLysLysCysAspGluLeuAlaAla 380  
Db 1081 AAGGGGGGAGACATCTCATCTTGTGTCATTCAAAGAAGAGTGGCAGAACTCGCCGCA 1140  
Qy 381 LysLeuValAlaLeuGlyValAsnAlaValAlaTyrTyrArgGlyLeuAspValSerVal 400  
Db 1141 AAGTGTGTGCTTGGGCATCAATCGCGTCTACTACCGCGTCTTGTGAGTCTCGGTC 1200  
Qy 401 IleProThrSerGlyAspValValValAlaThrAspAlaLeuMetThrGlyPheThr 420  
Db 1201 ATCCGCGCCATCGCGCATGTGTGTGTCGCAACCGATGCCCTCATGACCGGTATACC 1260  
Qy 421 GlyAspPheAspSerValIleAspCysAsnThrCysValThrGlnThrValAspPheSer 440

Db 1261 GGCAGCTTCGACTCGGTGATAGACTGCAATACAGTGTGTACCCAGACAGTGCATTCAGC 1320  
Qy 441 LeuAspProThrPheThrIleGluThrIleThrLeuProGlnAspAlaValSerArgThr 460  
Db 1321 CTTGACCTTACCTTACCATTTGACAAATCAGCTCCCCAAGATGCTGTCTCCGCACT 1380  
Qy 461 GlnArgArgGlyArgThrGlyArgGlyLysProGlyIleTyrArgPheValAlaProGly 480  
Db 1381 CAACGTCGGGGCAGGACTGGCAGGGGAGCCAGGCATCTACAGATTTGTGGCACCGGG 1440  
Qy 481 GluArgProSerGlyMetPheAspSerSerValLeuCysGluCysTyrAspAlaGlyCys 500  
Db 1441 GAGCGCCCTCCCGCATGTTTCGACTCGTCCGCTCTCTGTGAGTGTATGACCGAGGTGT 1500  
Qy 501 AlaTrpTyrGluLeuThrProAlaGluThrThrValArgLeuArgAlaTyrMetAsnThr 520  
Db 1501 GCTTGGTATGAGCTACGCCCCGCGAGACTACAGTTAGGCTTACGAGCGTACATGAACACC 1560  
Qy 521 ProGlyLeuProValCysGlnAspHisLeuGluPheTyrGluGlyValPheThrGlyLeu 540  
Db 1561 CCGGGGCTTCCCGTGTGCCAGGACCATTTGAATTTGGGAGGGGCTTTTACAGGCTC 1620  
Qy 541 ThrHisIleAspAlaHisPheLeuSerGlnThrLysGlnSerGlyGluAsnLeuProTyr 560  
Db 1621 ACTCATATAGATGCCACTTTCTATCCAGACAAAGCAGAGTGGGGAGAACCTTCTTAC 1680  
Qy 561 LeuValAlaTyrGlnAlaThrValCysAlaArgAlaGlnAlaProProProSerTrpAsp 580  
Db 1681 CTGGTAGCGTACCAGCCACCGTGTGCGTGTAGGGCTCAAGCCCTCCCCCATCGTGGAC 1740  
Qy 581 GlnMetTrpLysCysLeuIleArgLeuLysProThrLeuHisGlyProThrProLeuLeu 600  
Db 1741 CAGATGTGGAGTGTGTGATTGCTTCAAGCCACCTCCATGSGGCCAACACCTCTGCTA 1800  
Qy 601 TyrArgLeuGlyAlaValGlnAsnGluValThrLeuThrHisProValThrLysTyrIle 620  
Db 1801 TACAGACTGGGCGTGTTCAGAAATCACCCTGACGACCCAGTCCACCAATACATC 1860  
Qy 621 MetThrCysMetSerAlaAspLeuValValThrSerThrTrpValLeuValGlyGly 640  
Db 1861 ATGACATGCAATGTGGCGGACCTGGAGGTGTCTCAGAGCACCTGGGTGCTGTGGCGGC 1920  
Qy 641 ValLeuAlaAlaLeuAlaTyrCysLeuSerThrGlyCysValValIleValGlyArg 660  
Db 1921 GTCTGCTGCTTGTGGCGGCTTGTGCTGTCAACAGGCTGCGGTGTCATAGTGGGAGG 1980  
Qy 661 IleValLeuSerGlyLysProAlaIleIleProAspArgGluValLeuTyrArgGluPhe 680  
Db 1981 GTGCTCTTGTCCGGAAAGCGGCAATCATACCTGACAGGGAAGTCTCTTACCGAGAGTTC 2040  
Qy 681 AspGluMetGluGluCys 686  
Db 2041 GATGAGATGGAAGAGTGC 2058

## RESULT 13

US-09-881-654-1  
; Sequence 1, Application US/09881654  
; Patent No. 6632601  
; GENERAL INFORMATION:  
; APPLICANT: CHIEN, David Y.  
; APPLICANT: ARCANGEL, Phillip  
; APPLICANT: TANDESKE, Laura  
; APPLICANT: GEORGE-NASCIEMENTO, Carlos  
; APPLICANT: COLT, Doris  
; APPLICANT: MEDINA-SELBY, Angelica  
; TITLE OF INVENTION: IMMUNOASSAYS FOR ANTI-HCV ANTIBODIES  
; FILE REFERENCE: 2302-17039 / PP17039.002  
; CURRENT APPLICATION NUMBER: US/09/881,654  
; CURRENT FILING DATE: 2001-06-14  
; PRIOR APPLICATION NUMBER: 60/212,082  
; PRIOR FILING DATE: 2000-06-15  
; PRIOR APPLICATION NUMBER: 60/280,811

```
; PRIOR FILING DATE: 2001-04-02
; PRIOR APPLICATION NUMBER: 60/280,867
; PRIOR FILING DATE: 2001-04-02
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 1
; LENGTH: 2058
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:
; OTHER INFORMATION: representative NS3/4a conformational antigen
; NAME/KEY: CDS
; LOCATION: (1)..(2058)
US-09-881-654-1

Alignment Scores:
Pred. No.: 0 Length: 2058
Score: 3565.00 Matches: 671
Percent Similarity: 99.13% Conservative: 9
Best Local Similarity: 97.81% Mismatches: 6
Query Match: 98.54% Indels: 0
DB: 4 Gaps: 0

US-09-930-591-2 (1-686) x US-09-881-654-1 (1-2058)

QY 1 MetAlaProIleThrAlaTyrAlaGlnThrArgGlyLeuLeuGlyCysIleIleThr 20
DB 1 ATGGCGCCCATCAGCGGTACGCCACAGACAGCAAGGGGCTCTTAGGGTGCATAATCACC 60
QY 21 SerLeuThrGlyArgAspLysAsnGlnValGluGlyValGlnIleValSerThrAla 40
DB 61 AGCTTAATCGCGGGACAAAACCAAGTGGAGGGTGAGTCCAGATGTGTCACTGCT 120
QY 41 AlaGlnThrPheLeuAlaThrCysIleAsnGlyValCysTrpThrValTyrHisGlyAla 60
DB 121 GCCCAAACTTCTCGCAACGTCATCAATGGGGTGTGTGGACTGTGTACCAACGCGGCC 180
QY 61 GlyThrArgThrIleAlaSerProLysGlyProValIleGlnMetTyrThrAsnValAsp 80
DB 181 GGAACGAGGACCATCGCGTCACCAAGGGTCCGTGTCAATCAGATGTATACCAATGTAGAC 240
QY 81 GlnAspLeuValGlyTrpProAlaProGlnGlyAlaArgSerLeuThrProCysThrCys 100
DB 241 CAAGACCTTGTGGCTGGCCGCTCCGCAAGTAGCCGATCATGACACCTCGACTGC 300
QY 101 GlySerSerAspLeuTyrLeuValThrArgHisAlaAspValIleProValArgArgArg 120
DB 301 GGCTCCTCGGACCTTTACCTGGTCACGAGGACGCCGATGTCAATCCCGTCCGCGCGG 360
QY 121 GlyAspGlyArgGlySerLeuLeuSerProArgProIleSerTyrLeuLysGlySerSer 140
DB 361 GGTGATAGAGGGACGCCGTGTGTCGCCCGGCCCATTTCTTACTTTGAAGGCTCTCTCG 420
QY 141 GlyGlyProLeuLeuCysProAlaGlyHisAlaValGlyIlePheArgAlaAlaValCys 160
DB 421 GGGGTCCGTGTGTGTCGCCCGGGGACGCCGTGGGCATATTTAGGGCGCGGTGTGC 480
QY 161 ThrArgGlyValAlaLysAlaValAspPheIleProValGluSerLeuGluThrThrMet 180
DB 481 ACCCGTGGAGTGGCTAAGCGGTGCACTTTATCCCTGTGGAGAACCTTAGAGACAACCATG 540
QY 181 ArgSerProValPheSerAspAsnSerSerProProAlaValProGlnSerTyrGlnVal 200
DB 541 AGGTCCCGGTGTTCACGGATAACTCTCTCCACAGTAGTGCCCAAGAGCTTCCAGGTG 600
QY 201 AlahisLeuHisAlaProThrGlySerGlyLysSerThrLysValProAlaAlaTyrAla 220
DB 601 GCTCACCTCCATGCTCCACAGGACAGCGGCAAAAGCACCAAGGTCCCGGTGCATATGCA 660
QY 221 AlaGlnGlyTyrLysValLeuValLeuAsnProSerValAlaAlaThrMetGlyPheGly 240
DB 661 GCTCAGGGCTATAAGGTGCTAGTACTCAACCCCTCTGTGTGTGCAACACTGGGGCTTGGT 720
```

```
QY 241 AlaTyrMetSerLysAlaHisGlyIleAspProAsnIleArgThrGlyValArgThrIle 260
DB 721 GCTTACATGTCAGAGCTCATGGATCGATCTTAACATCAGGACCGGGGTGAGAACAAAT 780
QY 261 ThrThrGlySerProIleThrTyrSerThrTyrGlyLysPheLeuAlaAspGlyGlyCys 280
DB 781 ACCACTGGCAGGCCCATCATCAGTACTCACCTACGCAAGTTCCTTCCGACGCGCGGTGC 840
QY 281 SerGlyGlyAlaTyrAspIleIleIleCysAspGluCysHisSerThrAspAlaThrSer 300
DB 841 TCGGGGGGCGCTTATGACATAATAATTTGTGACGAGTGCACCTCCACGAGATGCCATCC 900
QY 301 IleLeuGlyIleGlyThrValLeuAspGlnAlaGluThrAlaGlyAlaArgLeuThrVal 320
DB 901 ATCTTGGGCATTTGGCACCTCTCTTGACCAAGCAGAGACTCGGGGGCGAGACTGGTGTG 960
QY 321 LeuAlaThrAlaThrProProGlySerValThrValProHisProAsnIleGluGluVal 340
DB 961 CTCGCCACGCCACACCTCGGGCTCCGTCACTGTGCCCATCCCAACATCGAGGAGGT 1020
QY 341 AlaLeuSerThrThrGlyGluIleProPheTyrGlyLysAlaIleProLeuGluAlaIle 360
DB 1021 GCTCTGTCCACCACCGAGAGATCCCTTTTACGGCAGGCTATCCCTCTCGAAGTAATC 1080
QY 361 LysGlyGlyArgHisLeuIlePheCysHisSerLysLysCysAspGluLeuAlaAla 380
DB 1081 AAGGGGGGAGACATCTCATCTCTGTCAATCAAAGAAGAGTGCAGCAACTCGCGCA 1140
QY 381 LysLeuValAlaLeuGlyValAsnAlaValAlaTyrTyrArgGlyLeuAspValSerVal 400
DB 1141 AAGCTGTCGCATTTGGGCATCAATGCGTGGCTACTACCGCGTCTTGACGTGTCGTC 1200
QY 401 IleProThrSerGlyAspValValValAlaThrAspAlaLeuMetThrGlyPheThr 420
DB 1201 ATCCGCCCATCGGCATGTGTGTCGTGGCAACCGATGCCCTCAAGCCGCTATACG 1260
QY 421 GlyAspPheAspSerValIleAspCysAsnThrCysValThrGlnThrValAspPheSer 440
DB 1261 GCGGACTTCGACTCGGTGATAGACTGCAATACGTGTGTCAACCCAGACAGTCGATTT 1320
QY 441 LeuAspProThrPheThrIleGluThrIleThrLeuProGlnAspAlaValSerArgThr 460
DB 1321 CTTGACCTTACCTTCACTTCCATTCAGACAATCAGCTCCCAAGATGCTGTCTCCGCACT 1380
QY 461 GlnArgArgGlyArgThrGlyArgGlyLysProGlyIleTyrArgPheValAlaProGly 480
DB 1381 CAACGTGGGGGACGAGTGGCAGGGGAGCCAGGCACTCTACAGATTTGTGGCACCGGG 1440
QY 481 GluArgProSerGlyMetPheAspSerSerValLeuCysGluCysTyrAspAlaGlyCys 500
DB 1441 GAGCGCCCTCCGCGCATGTTCGACTCGTCCGTCTCTGTGAGTGTATGACGCGGCTGT 1500
QY 501 AlaTyrTrpGluLeuThrProAlaGluThrThrValArgLeuArgAlaTyrMetAsnThr 520
DB 1501 GCTTGTATGAGTCAACGCCCGCGAGACTACAGTTAGGCTACGAGCGTACATGAACACC 1560
QY 521 ProGlyLeuProValCysGlnAspHisLeuGluPheTrpGluGlyValPheThrGlyLeu 540
DB 1561 CCGGGGCTTCCCGGTGGCCAGGACCATTTGAATTTTGGAGGGCGTCTTTACAGGCTC 1620
QY 541 ThrHisIleAspAlaHisPheLeuSerGlnThrLysGlnSerGlyGluAsnLeuProTyr 560
DB 1621 ACTCATATAGTCCCATTTTCTATCCAGACAAGCAGAGTGGGAGAACCTTCCTTAC 1680
QY 561 LeuValAlaTyrGlnAlaThrValCysAlaArgAlaGlnAlaProProSerTrpAsp 580
DB 1681 CTGGTAGCGTACCAGCCACCGGTGTGCTAGGGCTCAAGGCCCTCCCCCATCGTGGAC 1740
QY 581 GlnMetTrpLysCysLeuIleArgLeuLysProThrLeuHisGlyProThrProLeuLeu 600
DB 1741 CAGATGTGGAAGTGTGTGATTCGCTCAAGCCCAACCTCTCATGGGCCAACACCCCTGCTA 1800
```



QY 601 TyrArgLeuGlyAlaValGlnAsnGluValThrLeuThrHisProValThrIleValTyr 620  
Db 1801 TACAGACTGGCGCTGTTTTCAGAAATGAAATACCTTGCAGCGACCCAGTCCACCAATATAC 1860  
QY 621 MetThrCysMetSerAlaAspLeuGluValValThrSerThrTrpValLeuValGlyGly 640  
Db 1861 ATGACATGATGTCGGCGCGATCTGGAGGTGTCACGAGACCTGGTGCTCTGTTGGCGGC 1920  
QY 641 ValLeuAlaLeuAlaAlaTyrCysLeuSerThrGlyCysValValIleValGlyArg 660  
Db 1921 GTCTGGCTGCTTTGGCGCGCTATTGCTGCTCAACAGCGCTGGTGTATAGTGGCGAG 1980  
QY 661 IleValLeuSerGlyLysProAlaIleIleProAspArgGluValLeuTyrArgGluPhe 680  
Db 1981 GTGCTGTTGTCGGGAAGCGGCAATACATACCTGACAGGAAGTCTCTACCGAGAGTTC 2040  
QY 681 AspGluMetGluGluCys 686  
Db 2041 CATGATGGAGAGTGC 2058

## RESULT 14

US-10-637-323-1  
; Sequence 1, Application US/10637323  
; Patent No. 6797809  
; GENERAL INFORMATION:  
; APPLICANT: CHIEN, David Y.  
; APPLICANT: ARANGEL, Phillip  
; APPLICANT: TANDESKE, Laura  
; APPLICANT: GEORGE-NASCIEMENTO, Carlos  
; APPLICANT: COLT, Doris  
; APPLICANT: MEDINA-SELBY, Angelica  
; TITLE OF INVENTION: IMMUNOASSAYS FOR ANTI-HCV ANTIBODIES  
; FILE REFERENCE: 2302-17039 / PP17039.002  
; CURRENT APPLICATION NUMBER: US/10/637,323  
; CURRENT FILING DATE: 2003-08-08  
; PRIOR APPLICATION NUMBER: US/09/881,654  
; PRIOR FILING DATE: 2001-06-14  
; PRIOR APPLICATION NUMBER: 60/212,082  
; PRIOR FILING DATE: 2000-06-15  
; PRIOR APPLICATION NUMBER: 60/280,811  
; PRIOR FILING DATE: 2001-04-02  
; PRIOR APPLICATION NUMBER: 60/280,867  
; PRIOR FILING DATE: 2001-04-02  
; NUMBER OF SEQ ID NOS: 7  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 1  
; LENGTH: 2058  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence:  
; OTHER INFORMATION: representative NS3/4a conformational antigen

FEATU

NAME/KEY: CDS

LOCATION: (1)..(2058)

US-10-637-323-1

Alignment Scores:

Pred. No.: 0 Length: 2058  
Score: 3565.00 Matches: 671  
Percent Similarity: 99.13% Conservative: 9  
Best Local Similarity: 97.81% Mismatches: 6  
Query Match: 98.54% Indels: 0  
DB: 4 Gaps: 0

US-09-930-591-2 (1-686) x US-10-637-323-1 (1-2058)

QY 1 MetAlaProIleThrAlaTyrAlaGlnGlnThrArgGlyLeuGlyCysIleIleThr 20  
Db 1 ATGGCGCCCATCAGCGGTAGCCGACAGACAAAGGGGCTCTTAGGGTGCATATACCC 60  
QY 21 SerLeuThrGlyArgAspLysAsnGlnValGluGlyGluValGlnIleValSerAla 40

Db 61 AGCCTAACTGGCCCGGCAAAAACCAAGTGAGGGTGAGGTCCAGATTGTGTCAACTGCT 120  
QY 41 AlaGlnThrPheLeuAlaThrCysIleAsnGlyValCysTrpThrValTyrHisGlyAla 60  
Db 121 GCCCAAACTTCTTGGCAACGTCATCAATGGGGTGTGCTGACTGTCTACACCGGGGCC 180  
QY 61 GlyThrArgThrIleAlaSerProLysGlyProValIleGlnMetTyrThrAsnValAsp 80  
Db 181 GGAACGAGGACCATCGGCTCACCCCAAGGTCCTGTCTATCCAGATGTATACCAATGTAGAC 240  
QY 81 GlnAspLeuValGlyTrpProAlaProGlnGlyAlaArgSerLeuThrProCysThrCys 100  
Db 241 CAAGACCTTGGGCTGGCCGCTCCGCAAGTAGCCGATCATTTGACACCTCGACCTGTC 300  
QY 101 GlySerSerAspLeuTyrLeuValThrArgHisAlaAspValIleProValArgArg 120  
Db 301 GGCTCCTCGGACCTTTACCTGGTCAAGGAGCAGCCGATGTCTATCCGTCGGCGGGCGG 360  
QY 121 GlyAspGlyArgGlySerLeuLeuSerProArgProIleSerTyrLeuLysGlySerSer 140  
Db 361 GGTGATAGAGGGGACGCTGCTGTCGCCCGGCCCATTTCTACTTTGAAAGGCTCCTCG 420  
QY 141 GlyGlyProLeuLeuCysProAlaGlyHisAlaValGlyIlePheArgAlaAlaValCys 160  
Db 421 GGGGTCCTGCTGTTGTCGCCCGGGGACGCGCTGGGCATATTTAGGGCGCGGCTGTGC 480  
QY 161 ThrArgGlyValAlaLysAlaValAspPheIleProValGluSerLeuGluThrMet 180  
Db 481 ACCGTGGAGTGGCTAAGCGGTGGACTTTATCCTGTGGAGAACCTTAGAGAACCAATG 540  
QY 181 ArgSerProValPheSerAspAsnSerSerProProAlaValProGlnSerTyrGlnVal 200  
Db 541 AGGTCCCGGTGTTACGGGATAACTCTCTCCACCACTAGTGCCTCCAGAGCTTCAGGTG 600  
QY 201 AlaHisLeuHisAlaProThrGlySerGlyLysSerThrLysValProAlaAlaTyrAla 220  
Db 601 GCTCACCTCCATGCTCCACAGGCGAGGCAAAAGCACCAAGGTCCGCGGTGCATATGCA 660  
QY 221 AlaGlnGlyTyrLysValLeuValLeuAsnProSerValAlaAlaThrMetGlyPheGly 240  
Db 661 GCTCAGGGCTATAGGTGCTAGTACTCAACCCCTCTGTGTGCAACACTGGGCTTTGGT 720  
QY 241 AlaTyrMetSerLysAlaHisGlyIleAspProAsnIleArgThrGlyValArgThrIle 260  
Db 721 GCTTACATGTCGAAGGCTCATGGATCGATCTTAACATCAGGACCGGGGTGAGAACAT 780  
QY 261 ThrThrGlySerProIleThrTyrSerThrTyrGlyLysPheLeuAlaAspGlyCys 280  
Db 781 ACCACTGGCAGCCCATCAGGTACTCCACCTAGCGCAAGTTCTTTCGCGCGGGGTGC 840  
QY 281 SerGlyGlyAlaTyrAspIleIleIleCysAspGluCysHisSerThrAspAlaThrSer 300  
Db 841 TCGGGGGCGCTTATGACATAAATTTGTGAGAGTGCCACTCCACGATGCCACATCC 900  
QY 301 IleLeuGlyIleGlyThrValLeuAspGlnAlaGluThrAlaGlyAlaArgLeuThrVal 320  
Db 901 ATCTTGGCAFTGGCACTGCTTGGACAGCAGAGACTCGGGGGGAGACTGGTGTG 960  
QY 321 LeuAlaThrAlaThrProProGlySerValThrValProHisProAsnIleGluGluVal 340  
Db 961 CTGCGCACCGCCACCCCTCCGGGCTCGTCACTGTGCCCCCATCCCAACATCGAGGAGTT 1020  
QY 341 AlaLeuSerThrThrGlyGluIleProPheTyrGlyLysAlaIleProLeuGluAlaIle 360  
Db 1021 GCTCTGTCCACCGGAGAGATCCCTTTTACGGCAAGGCTATCCCTCCGAAAGTAATC 1080  
QY 361 LysGlyGlyVArgHisLeuIlePheCysHisSerLysLysCysAspGluLeuAlaAla 380  
Db 1081 AAGGGGGGAGAGATCTCATCTCTTGTCAATCAAAAGAAAGTGCAGCAACTCGCCGCA 1140  
QY 381 LysLeuValAlaLeuGlyValAsnAlaValAlaTyrTyrArgGlyLeuAspValSerVal 400  
Db 1141 AAGCTGCTCGCATTTGGGCATCAATCGCTGCGCTACTACCGGGTCTTTGACGTGCTCGTC 1200



QY 401 IleProThrSerGlyAspValValValAlaThrAspAlaLeuMetThrGlyPheThr 420  
Db 1201 ATCCGCCCATCGCGATGTTGCTGCTGCAACCGATGCCCTCATGCCGCTATACG 1260  
QY 421 GlyAspPheAspSerValIleAspCysAsnThrCysValThrGlnThrValAspPheSer 440  
Db 1261 GCGGACTTCGACTCGGTGATAGACTGCAATACCTGTGTACCCAGACAGCTGATTTTCAGC 1320  
QY 441 LeuAspProThrPheThrIleGluThrIleThrLeuProGlnAspAlaValSerArgThr 460  
Db 1321 CTGACCTTACCTTCCATTGACATGACATACATCCGCTCCCCAGATGCTGTCTCCGCACT 1380  
QY 461 GlnArgArgGlyArgThrGlyArgGlyLysProGlyIleTyrArgPheValAlaProGly 480  
Db 1381 CAACGTCGGGCGAGGACTGCGCGGGAAGCCAGGCACTACAGATTTGTGGCACCGGG 1440  
QY 481 GluArgProSerGlyMetPheAspSerSerValLeuCysGluCysTyrAspAlaGlyCys 500  
Db 1441 GAGCGCCCTCCGCGCATGTTTCGACTCGTCCGCTCTCTGTGAGTGTATGACGCGAGGTGT 1500  
QY 501 AlaTyrTyrGluLeuThrProAlaGluThrThrValArgLeuArgAlaTyrMetAsnThr 520  
Db 1501 GCTTGGTATGAGTTCAGCGCCGCGAGACTACAGTTAGGCTACGAGCGTACATGAACACC 1560  
QY 521 ProGlyLeuProValCysGlnAspHisLeuGluPheTyrGluGlyValPheThrGlyLeu 540  
Db 1561 CCGGGCTTCCTCGTGTGCCAGGACCATCTTGAATTTTGGGAGGCGCTTTTACAGGCGCTC 1620  
QY 541 ThrHisIleAspAlaHisPheLeuSerGlnThrLysGlnSerGlyGluAsnLeuProTyr 560  
Db 1621 ACTCATATAGATCCCACTTCTATCCAGACAAAGCAGAGTGGGAGAACCTTCTCTTAC 1680  
QY 561 LeuValAlaTyrGlnAlaThrValCysAlaArgAlaGlnAlaProProSerTyrAsp 580  
Db 1681 CTGGTACGTAACCAAGCCCGTGTGGCTAGGCTCAAGCCCTCCCTCCCTCCCTCGGGAC 1740  
QY 581 GlnMetTyrLysCysLeuIleArgLeuLysProThrLeuHisGlyProThrProLeuLeu 600  
Db 1741 CAGATGTGGAGTGTGTGATGCTTCAAGCCCAACCTCATGTCGCAACACCCCTGCTA 1800  
QY 601 TyrArgLeuGlyAlaValGlnAsnGluValThrLeuThrHisProValThrLysTyrIle 620  
Db 1801 TACAGACTGGCGCTGTTCAGATGAATCAATCAGCTCAGCAGCAGCAGTCAACAAATACATC 1860  
QY 621 MetThrCysMetSerAlaAspLeuGluValValThrSerThrTrpValLeuValGly 640  
Db 1861 ATGACATGCAATGTCGGCGGACCTGGAGTGTCTCAGAGCACTGGGTGCTGTGGCGGC 1920  
QY 641 ValLeuAlaAlaLeuAlaTyrCysLeuSerThrGlyCysValValIleValGlyArg 660  
Db 1921 GTCTGGCTGCTTTGGCGCGTATTCCTCTCAACAGGCTCGTGTGATAGTGGCAGG 1980  
QY 661 IleValLeuSerGlyLysProAlaIleIleProAspArgGluValLeuTyrArgGluPhe 680  
Db 1981 GTGCTGTGTCCGGAAGCGGCAATCATACCTGACAGGGAAGTCTCTCTACCGAGATTC 2040  
QY 681 AspGluMetGluGluCys 686  
Db 2041 GATGAGATGAAGAGTGC 2058

RESULT 15  
US-08-444-818-137  
; Sequence 137, Application US/08444818  
; Patent No. 6150087  
; GENERAL INFORMATION:  
; APPLICANT: Chien, David Y.  
; APPLICANT: Rutter, William J.  
; TITLE OF INVENTION: NANV Diagnostics and Vaccines  
; NUMBER OF SEQUENCES: 777  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Chiron Corporation  
; STREET: 4560 Horton Street

CITY: Emeryville  
STATE: CA  
COUNTRY: USA  
ZIP: 94608-2916  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/444,818  
FILING DATE:  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/403,590  
FILING DATE: 14-MAR-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Harbin, Alisa A.  
REGISTRATION NUMBER: 33,895  
REFERENCE/DOCKET NUMBER: 0110.002  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (508)359-3876  
TELEFAX: (508)359-3885  
INFORMATION FOR SEQ ID NO: 137:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 8987 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 1..8985  
US-08-444-818-137

Alignment Scores:  
Pred. No.: 0 Length: 8987  
Score: 3565.00 Matches: 671  
Percent Similarity: 99.27% Conservative: 10  
Best Local Similarity: 97.81% Mismatches: 5  
Query Match: 98.54% Indels: 0  
DB: 3 Gaps: 0

US-09-930-591-2 (1-686) x US-08-444-818-137 (1-8987)

QY 1 MetAlaProIleThrAlaTyrAlaGlnThrArgGlyLeuLeuGlyCysIleIleThr 20  
Db 3076 CTGGCGCCCATCATCGCGCTACGCCAGCAGACAGAGGGGCTCTTAGGGTGCATATCACC 3135  
QY 21 SerLeuThrGlyArgAspLysAsnGlnValGluGlyGluValGlnIleValSerThrAla 40  
Db 3136 AGCTTAACCTGGCGGACAAACCAAGTGGAGGTGAGGTCCAGATTGTGTCAACTGCT 3195  
QY 41 AlaGlnThrPheLeuAlaThrCysIleAsnGlyValCysTrpThrValTyrHisGlyAla 60  
Db 3196 GCCCAAACTTCTCGCAACGTGCATCAATGGGGTGTCTGGACTGTCTACCAACGGGGCC 3255  
QY 61 GlyThrArgThrIleAlaSerProLysGlyProValIleGlnMetTyrThrAsnValAsp 80  
Db 3256 GGAACGAGGACCATCGGTCACCAAGGTCTCTGTCTATCCAGATGTATACCATGTAGAC 3315  
QY 81 GlnAspLeuValGlyTrpProAlaProGlnGlyAlaArgSerLeuThrProCysThrCys 100  
Db 3316 CRAAGACCTTGTGGGCTGGCGCTCGCAAGGTAGCGGCTCATTCAGACCTCCTGCTGC 3375  
QY 101 GlySerSerAspLeuTyrLeuValThrArgHisAlaAspValIleProValArgArgArg 120  
Db 3376 GGCTCTCGGACCTTTTACCTGGTGCAGGACGACCGGATGTCTATCCCGTCCCGCGCGG 3435  
QY 121 GlyAspGlyArgGlySerLeuLeuSerProArgProIleSerTyrLeuLysGlySerSer 140  
Db 3436 GGATGATAGCGGGGACGCTGTCTGTGCGCCCGCCCATTTCTTACTTTGAANGSCTCTCG 3495

```
Qy 141 GlyGlyProLeuLeuCysProAlaGlyHisAlaValGlyIlePheArgAlaAlaValCys 160
Db 3496 GGGGGTCGGCTGGTGTGCGCGCGGGCACCGCGGGCATATTTAGGGCGCGGGTGTGC 3555
Qy 161 ThrArgGlyValAlaLysAlaValAspPheIleProValGluSerLeuGluThrThrMet 180
Db 3556 ACCCGTGGAGTGGCTAAGCGGGTGACCTTTATCTCTGTGGAGAACCTTAGAGACAAACCATG 3615
Qy 181 ArgSerProValPheSerAspAsnSerSerProProAlaValProGlnSerTyrGlnVal 200
Db 3616 AGGTCCCGGTGTTACAGGATAACTCTCTCCACCAGTAGTGCCCGCAGAGCTTCCAGGTG 3675
Qy 201 AlaHisLeuHisAlaProThrGlySerGlyLysSerThrLysValProAlaAlaTyrAla 220
Db 3676 GCTCACTCCATGCTCCACAGGAGCGGCAAGGACCAAGGTCGCGGCTGCATATGCA 3735
Qy 221 AlaGlnGlyTyrLysValLeuValLeuAsnProSerValAlaAlaThrMetGlyPheGly 240
Db 3736 GCTCAGGGCTATAGGTGCTAGTACTCAACCCCTCTGTGTCTGCAACTCGGGCTTTGGT 3795
Qy 241 AlaTyrMetSerLysAlaHisGlyIleAspProAsnIleArgThrGlyValArgThrIle 260
Db 3796 GCTTACATGTCCAGGGCTCATGGATCGATCCTTAACATCAGACCGGGTGAGAACAAAT 3855
Qy 261 ThrThrGlySerProIleThrTyrSerThrTyrGlyLysPheLeuAlaAspGlyGlyCys 280
Db 3856 ACCACTGGCAGCCCCATCACGTACTCCACCTACGCGCAAGTTCTTTGGCGAGCGGGGTGC 3915
Qy 281 SerGlyGlyAlaTyrAspIleIleCysAspGluCysHisSerThrAspAlaThrSer 300
Db 3916 TCGGGGGCGCTTATGACATTAATTTGTACAGAGTGCCACTCCACGGATGCCACATCC 3975
Qy 301 IleLeuGlyIleGlyThrValLeuAspGlnAlaGluThrAlaGlyAlaArgLeuThrVal 320
Db 3976 ATCTTGGGCATCGGACGTGCTTGACCAACGACGAGACTGCGGGGGCGAGACTGGTTGTG 4035
Qy 321 LeuAlaThrAlaThrProGlySerValThrValProHisProAsnIleGluGluVal 340
Db 4036 CTGCGCCACCGCACCCCTCCGGGTCCGGTCACTGTGCCCATCCCAACATCGAGAGGTT 4095
Qy 341 AlaLeuSerThrThrGlyGluIleProPheThrGlyLysAlaIleProLeuGluAlaIle 360
Db 4096 GCTCTGTCCACCCGAGAGATCCCTTTTACGGCAAGGCTATCCCCCTCGAAGTAATC 4155
Qy 361 LysGlyGlyArgHisLeuIlePheCysHisSerLysLysCysAspGluLeuAlaAla 380
Db 4156 AAGGGGGGAGACATCTCTCTGTCAATCAAGAAGAGTGGCGAGACTCGCCGCA 4215
Qy 381 LysLeuValAlaLeuGlyValAsnAlaValAlaTyrTyrArgGlyLeuAspValSerVal 400
Db 4216 AAGCTGGTTCGCTATGGGCATCAATGCGGTGGCTTACTACCGCGGTCTTGACGTGTCGTC 4275
Qy 401 IleProThrSerGlyAspValValValAlaThrAspAlaLeuMetThrGlyPheThr 420
Db 4276 ATCCCGACACGCGCGATGTGTGCTGTGGCAACCGATGCCCTCATGACGGCTATACC 4335
Qy 421 GlyAspPheAspSerValIleAspCysAsnThrCysValThrGlnThrValAspPheSer 440
Db 4336 GGCACCTCGACTCGGTAGATGTCATACGTGTGTACCCAGACAGTCGATTTTCCAGC 4395
Qy 441 LeuAspProThrPheThrIleGluThrIleThrLeuProGlnAspAlaValSerArgThr 460
Db 4396 CTTGACCTTACCTTCAACATTGAGACAATACGCTCCCGAGGATGCTGTCTCCGCACT 4455
Qy 461 GlnArgArgGlyArgThrGlyArgGlyIlePheProGlyIleTyrArgPheValAlaProGly 480
Db 4456 CAACGTCGGGCAAGACTGGCAGGGGGAAGCAGGCATCAACAGATTGTGGCACCGGGG 4515
Qy 481 GluArgProSerGlyMetPheAspSerSerValLeuCysGluCysTyrAspAlaGlyCys 500
Db 4516 GAGCGCCCTCCGGCATGTTGACCTCGTCCGTCTCTGTGAGCTATGACGAGGCTGT 4575
Qy 501 AlaTrpTyrGluLeuThrProAlaGluThrThrValArgLeuArgAlaTyrMetAsnThr 520
```

```
Db 4576 GCTTGGTATGAGCTCACGCCCGCGAGACTACAGTTAGGCTACGAGCGTACATGAACACC 4635
Qy 521 ProGlyLeuProValCysGlnAspHisLeuGluPheTrpGluGlyValPheThrGlyLeu 540
Db 4636 CCGGGGCTTCCCGTGTGCCAGGACCACTTGAATTTTGGGAGGGGCTTTTACAGGCCTC 4695
Qy 541 ThrHisIleAspAlaHisPheLeuSerGlnThrLysGlnSerGlyGluAsnLeuProTyr 560
Db 4696 ACTCATATAGATGCCACTTCTATCCAGACAAAGCAGAGTGGGGAGAACCTTCTTCTAC 4755
Qy 561 LeuValAlaTyrGlnAlaThrValCysAlaArgAlaGlnAlaProProProSerTrpAsp 580
Db 4756 CTGCTAGCGTACCAAGCACCGTGTGCGCTCAGGCTCAAGCCCTCCCCCATCGTGGGAC 4815
Qy 581 GlnMetTrpLysCysLeuIleArgLeuLysProThrLeuHisGlyProThrProLeuLeu 600
Db 4816 CAGATGTGGAAGTGTGTTGATTCGCTCAGCCCACTTCCATGGGCCCAACCCCTGTCTA 4875
Qy 601 TyrArgLeuGlyAlaValGlnAsnGluValThrLeuThrHisProValThrLysTyrIle 620
Db 4876 TACAGACTGGGCGCTGTTCAAGATGAATCACCCTGACGCAACCCAGTCACCAATATCATC 4935
Qy 621 MetThrCysMetSerAlaAspLeuGluValValThrSerThrTrpValLeuValGlyGly 640
Db 4936 ATGACATGTCATGTGCGCGCACCTGGAGGTGCTCAGAGCACCTGGGTGCTCGTTGGCGGC 4995
Qy 641 ValLeuAlaAlaLeuAlaAlaTyrCysLeuSerThrGlyCysValValIleValGlyArg 660
Db 4996 GTCTGTGCTGCTTTGGCGCGCTATTGCTGTCAACAGCGTGTGGTGTATAGTGGGAGG 5055
Qy 661 IleValLeuSerGlyLysProAlaIleIleProAspArgGluValLeuTyrArgGluPhe 680
Db 5056 GTGCTCTTGTCCGGAGAGCGGCAATCATACCTGACAGGGAGAGTCTCTACCGAGAGTTC 5115
Qy 681 AspGluMetGluGluCys 686
Db 5116 GATGAGATGGAAGAGTGC 5133
```

Search completed: February 27, 2005, 07:25:22  
Job time : 412 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: February 27, 2005, 03:26:17 ; Search time 4728 Seconds  
(without alignments)  
5522.861 Million cell updates/sec

Title: US-09-930-591-2

Perfect score: 3618

Sequence: 1 MAPTAYAQTRGLLGCIIIT.....PAIPDREVLVREDFEMEEC 686

Scoring table:

BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DEV=xlp  
-Q=/cgn2\_1/USPTO\_spool\_p/US930591/runat\_25022005\_143613\_25686/app\_query.fasta\_1.839  
-DB=EST -QFMT=fastap -SUFFIX=src -MINMATCH=0.1 -LOOPEL=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45  
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US09930591 @CGN 1.1 5180 @runat 25022005 143613 25686 -NCPU=6 -ICPU=3  
-NO.MMAP -LARGQUERY -NEG SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG  
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=10 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

EST:\*  
1: gb\_est1:\*  
2: gb\_est2:\*  
3: gb\_hcc:\*  
4: gb\_est3:\*  
5: gb\_est4:\*  
6: gb\_est5:\*  
7: gb\_est6:\*  
8: gb\_gse1:\*  
9: gb\_gse2:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Match	Query Length	ID	Description
1	181	5.0	2236	3	CR591941 full-leng
2	163	4.5	3291	3	BC046781 Mus muscu
3	160	4.4	3201	9	CL974652 OsIFCC026
4	150.5	4.2	3229	3	AK004665 Mus muscu
5	141	3.9	3359	9	AY411078 Homo sapi
6	140.5	3.9	822	5	BU054791 UI-M-FD0-
7	136.5	3.8	3320	3	BC059369 Homo sapi
8	133.5	3.7	889	6	CD359697 AGENCOURT
9	132.5	3.7	818	5	BU187274 AGENCOURT

10	131	3.6	949	1	AL560974
11	129.5	3.6	2856	3	CR613419 full-leng
12	129	3.6	2972	3	AK031534 Mus muscu
13	129	3.6	2984	3	AK031679 Mus muscu
14	129	3.6	2986	3	AK028274 Mus muscu
15	129	3.6	3956	3	BC043699 Mus muscu
16	128.5	3.6	1697	9	AY404177 Mus muscu
17	128	3.5	636	7	CN788030 4122309 B
18	128	3.5	845	7	CV106671 AGENCOURT
19	127.5	3.5	691	5	BU054966 UI-M-FD0-
20	127.5	3.5	738	6	CD240900 AGENCOURT
21	127.5	3.5	2976	9	AY400284 Homo sapi
22	126	3.5	1283	5	BQ709745 AGENCOURT
23	126	3.5	1642	3	CR610486 full-leng
24	125.5	3.5	2484	3	CNS0AAU3 Arabidops
25	124.5	3.4	623	7	CF131593 UI-HF-F00
26	124.5	3.4	2388	9	AY418898 Mus muscu
27	124.5	3.4	3633	3	AY383690 Rattus no
28	124	3.4	1185	9	CL962117 OsIFCC007
29	124	3.4	2388	9	AY418896 Homo sapi
30	123.5	3.4	1788	3	AY105041 Zea mays
31	123.5	3.4	1818	9	CL973139 OsIFCC042
32	123	3.4	790	7	CK017540 AGENCOURT
33	123	3.4	3461	9	AY398774 Mus muscu
34	123	3.4	3802	3	AK004733 Mus muscu
35	123	3.4	4327	3	BC058331 Mus muscu
36	123	3.4	4640	3	BC062885 Mus muscu
37	122	3.4	3879	3	BC028405 Homo sapi
38	121.5	3.4	3211	3	BC041392 Homo sapi
39	121	3.3	2270	9	AY418897 Pan trogl
40	121	3.3	7320	3	CR627021 Homo sapi
41	120.5	3.3	478	2	BF725559 bx18a02.y
42	120.5	3.3	694	4	BI088407 602852901
43	120.5	3.3	3856	3	AK078552 Mus muscu
44	120	3.3	651	1	AJ729569 AJ729569
45	120	3.3	1143	5	BM926541 AGENCOURT

#### ALIGNMENTS

RESULT 1  
CR591941  
LOCUS  
DEFINITION full-length cDNA clone CS0DL006YA04 of B cells (Ramos cell line)  
2236 bp mRNA linear HTC 21-JUL-2004  
CR591941  
Accession  
Version CR591941.1 GI:50472748  
Keywords HTC; CNSLT\_CDNA.  
Source Homo sapiens (human)  
Organism Homo sapiens  
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 2236)  
Li, W.B., Gruber, C., Jesses, J. and Polayes, D.  
Full-length cDNA libraries and normalization  
Unpublished  
Contact: Feng Liang Email: fliang@lifetech.com URL: http://fulllength.invitrogen.com/ Invitrogen Corporation 1600 Faraday Avenue  
2 (bases 1 to 2236)  
Genoscope.  
Direct Submission  
Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail: seque@genoscope.cns.fr)  
- Web : www.genoscope.cns.fr)  
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life technologies, a division of Invitrogen.  
Location/Qualifiers  
1. .2236  
/organism="Homo sapiens"



## AUTHORS

Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G., Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D., Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K., Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F., Ditchenko, L., Marusica, K., Farmer, A.A., Rubin, G.M., Hong, L., Stapleton, M., Soares, M.B., Bonaldo, M.P., Casavant, T.L., Scheetz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S., Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J., Abramson, R.D., Mullahy, S.J., Bosak, S.A., McEwan, P.J., McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S., Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W., Villalon, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Fahey, J., Helton, E., Kettner, M., Madan, A., Young, A.C., Rodriguez, S., Sanchez, A., Whiting, M., Madan, A., Touchman, J.W., Green, E.D., Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D., Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Butlerfield, Y.S., Krzyzanski, M.I., Skalek, U., Smalick, D.E., Schnerch, A., Schein, J.E., Jones, S.J., and Marra, M.A.

**TITLE**  
Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences

**JOURNAL**  
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

**MEDLINE**  
22388257

**PUBMED**  
12477932

**REFERENCE**  
2 (bases 1 to 3291)

**AUTHORS**  
Strausberg, R.

**TITLE**  
Direct Submission

## JOURNAL

Submitted (13-FEB-2003) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA

**REMARK**  
NIH-MGC Project URL: <http://mgc.nci.nih.gov>

**COMMENT**  
Contact: MGC help desk

Email: [cgapbs-remail.nih.gov](mailto:cgapbs-remail.nih.gov)

Tissue Procurement: Dr. Jim Lin, University of Iowa

cDNA Library Preparation: M. Bento Soares, University of Iowa

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)

DNA Sequencing by: University of Iowa, Dr. M. Bento Soares and Dr. Thomas L. Casavant

Web site: <http://genome.uiowa.edu>

Contact: [bento-soares@uiowa.edu](mailto:bento-soares@uiowa.edu); [tom-casavant@uiowa.edu](mailto:tom-casavant@uiowa.edu)

Bonaldo, M.F., Akabogu, I., Bair, T., Bair, J., Crouch, K., Davis, A., Fishler, K., Keppel, C., Kucaba, T., Lebeck, M., Melo, A., Schaefer, K., Scheetz, T., Smith, C., Snir, E., Tack, D., Trout, K., Walters, J., Casavant, T., Soares, M.B.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: <http://image.llnl.gov>

Series: Plate: Row: Column: 0

This clone was selected for full length sequencing because it passed the following selection criteria: Similarity but not identity to protein

This clone has the following problem: frame shifted.

Location/Qualifiers

1. 3291

/organism="Mus musculus"

/mol\_type="mRNA"

/strain="C57BL/6"

/db\_xref="taxon:10090"

/clones="IMAGE:5716182"

/tissue\_type="Mouse, brain 12.5 dpc"

/clone\_lib="NTH BWAP\_FCO"

/lab\_host="DH10B"

/note="Vector: pYX-ASC"

## ORIGIN

Alignment Scores:

Pred. No.:	0.000876	Length:	3291
Score:	163.00	Matches:	127
Percent Similarity:	35.15%	Conservative:	60
Best Local Similarity:	23.87%	Mismatches:	165
Query Match:	4.51%	Indels:	181
DB:	3	Gaps:	27

US-09-930-591-2 (1-686) x BC046781 (1-3291)

Qy 89 ProGlnGlyAlaArgSerLeuThrPro-----CysThrCys----- 100  
 |||||  
 |||||  
 Db 953 CCGGAGACGAGGAGAGCTGGAGGCCACTAATCGTACCACATGCCCAAGGAGAGCCGAG 1012  
 |||||  
 |||||  
 Qy 101 GlySerSer---AspLeuThrValThrArgHisAlaAspValIleProValArgArg 119  
 |||||  
 |||||  
 Db 1013 GCGAGCCAGCTCGAACTGTGACATCGTTGAGGAGGAATCGGTCGCGGAGAGGAGC 1072  
 |||||  
 |||||  
 Qy 120 ArgGlyAspGlyArg-Gly-----SerLeuLeuSerProArgPr 132  
 |||||  
 |||||  
 Db 1073 AGCGCGCTGGGAGAGGCCCGCAGCTGGGTGTGTGATCCCTGAAGTTGGAGCCGAGATG 1132  
 |||||  
 |||||  
 Qy 132 oileSerTyrLeuLysGlySerSerGlyProLeuLeuCysProAlaGly----- 149  
 |||||  
 |||||  
 Db 1133 CTGC-----TCGCGAGGAGGCCAAGTACCAGCTGGTGTGGAGGAGGAGC 1177  
 |||||  
 |||||  
 Qy 150 ----HisAlaVal-----GlyIlePheArgAlaAl 158  
 |||||  
 |||||  
 Db 1178 AGACCATCGAGTTGTCCGTGCTCAACTCCAGGGTGACGAGGAGCGCTCAGGCCGC 1237  
 |||||  
 |||||  
 Qy 158 aValCysThrArgGlyValAla-LysAlaValAspPheIleProValGluSerLeuGlu 178  
 |||||  
 |||||  
 Db 1238 CCCTGTCCAGCCCGAGCCGAGCAGAG-----GAATCTATCCAGG 1276  
 |||||  
 |||||  
 Qy 178 hrThrMetArgSer---ProValPheSerAspAsnSerSerProProAlaValProGlns 197  
 |||||  
 |||||  
 Db 1277 CTGTGCGCGCAGCTGCGCGTGTCCCTCCGAGGAGAGCTTCTGCCCGCATTGCCA 1336  
 |||||  
 |||||  
 Qy 197 erTyrGlnValAlaHisLeuHisAlaProThrGlySerGlyLysSerThrLysValPro- 216  
 |||||  
 |||||  
 Db 1337 ACCATCAGTCTCATCATCGAAGCGGAGAGCTGCTCTGGGAGACACACAGATCCAC 1396  
 |||||  
 |||||  
 Qy 217 -----AlaAlaTyrAlaAlaGlnGlyTyrLysValLeuValLeuAsnProS 232  
 |||||  
 |||||  
 Db 1397 AGTACCTCTTTTGGAGGAGGTTACACAAAGAGGCGATGAGATGCTTGCACCCGCC 1456  
 |||||  
 |||||  
 Qy 232 erValAlaAlaThrMetGlyPheGlyAlaTyrMetSerLysAlaHisGly----- 248  
 |||||  
 |||||  
 Db 1457 GGAGAGTGGCGGTATGAGTGTGCGAGCCCGAGTGGCGCGGAGATGGTGTGAAGCTTG 1516  
 |||||  
 |||||  
 Qy 249 -----IleAspProAsnIleArg-----ThrGlyValArgThrIleThr 262  
 |||||  
 |||||  
 Db 1517 GGAACGAGGTGGGTACAGCATCCGTTGGAGGACTGCACCTCAGAGGCACTGT----- 1572  
 |||||  
 |||||  
 Qy 262 hrGlySerProIleThrTyrSerThrTyrGly-----LysPheLeuAlaAspG 278  
 |||||  
 |||||  
 Db 1573 -----CTCCGCTACATGACAGATGGAATGCTACTCCGAGAGTCTCTCTGAGC 1621  
 |||||  
 |||||  
 Qy 278 lyGlyCysSerGlyValAlaTyrAspIlelleCysAspGluCysHisSerThrAspA 298  
 |||||  
 |||||  
 Db 1622 CTGACCTTGCA-----AGTTACAGTGTGGTGTGGATGAAGCTCAGAGCGGACCT 1675  
 |||||  
 |||||  
 Qy 298 laThrSerIleLeuGlyIleGlyThrValLeuAspGlnAlaGluThrAlaGlyAlaArgL 318  
 |||||  
 |||||  
 Db 1676 TGCACACAGACATCTCTTTGGATGATCAAGACGCTGCTAGATTCGACCTGAGCTCA 1735  
 |||||  
 |||||  
 Qy 318 euThrValLeuAlaThrAlaThrProProGlySerValThrValProHisProAsnIleG 338  
 |||||  
 |||||  
 Db 1736 AGGTCTGTGGTTCAGCCACA----- 1758  
 |||||  
 |||||  
 Qy 338 luGluValAlaLeuSerThrThrGlyGluIleProPheTyrGlyLysAlaIleProLeuG 358  
 |||||  
 |||||  
 Db 1759 -----CTGGATACCTGCCCGTCTTCTGCTCTTCTCGATGAGGCC---CCTGTCT 1804  
 |||||  
 |||||  
 Qy 358 luAlaIleLysGlyGlyArgHisLeu-----IlePheCysHisSerLysLysCysA 376  
 |||||  
 |||||  
 Db 1805 TCAGAAATCCCTGGACGAGGTTTCCAGTTGACATCTTCTATACCAGGCCCCAGAGGCTG 1864  
 |||||  
 |||||  
 Qy 376 spGluLeuAlaAlaLysLeuValAlaLeuGlyValAsnAlaValAlaTyrTyrArgGlyL 396  
 |||||  
 |||||  
 Db 1865 ACTACCTGGAAGCTCGCTGTGTCTGTG----- 1894

```

QY 396 euAspValSerValIleProThrSerGlyAspValValValValAlaThrAspAlaLeuM 416
Db 1895 TGCAGATCCACGTCACCCCGCTGGAGATATATCTGGTG-----TTCC 1939
QY 416 etThrGlyPheThrGlyAspPheAspSerValIleAspCys-----429
Db 1940 TGACGGGACA-GGAGGAGATTGAGGCTGCTGTGAGATGCTCCAGGACCGCTCCGCGAG 1998
QY 429 -----429
Db 1999 CTGGGCTCCAAGATCCGGGAGCTCTGGNGCTGCCCATTTATGCCAACCTGCCCTCAGAC 2058
QY 430 -----AsnThrCysValThr- 434
Db 2059 ATGCAGGCTCGCATCTTCCAGCCCAACACCCCGGGGCCGGAAGAACTTCCCTCACCA 2118
QY 435 --GlnThrValAspPheSerLeuAspProThrPheThr-----446
Db 2119 TCGAAGGCATCATCTATGTCTGGACCCAGGGTTCTGCAAGCAGAGAGCTACAACCCCTC 2178
QY 447 -----IleGluThrIleThrLeu---ProGlnAspAlaValSerArgThrGlnArgA 463
Db 2179 GTACGGGAATGGAGTCACTCACCGTCAACCCCTGCAGCAAGGCTTCAGCCCAATCAGCGGG 2238
QY 463 rgGlyArgThrGlyArgGlyLysProGlyIleTyArgPheValAlaProGlyGluArgP 483
Db 2239 CTGGCCCTGCAGTCA-----GTGGCTGCCGG-----2267
QY 483 roSerGlyMetPheAspSerSerValLeuCysGluCysTyArgAspAlaGlyCysAlaTrp- 502
Db 2268 -----AAGTCTTCGCGCTGTATACGGCCTGGG 2295
QY 503 --TyrGluLeuThrProAlaGluThrThrVal 512
Db 2296 CCTATCAGCATGCTAGAGGAGACACAGATT 2327

RESULT 3
LOCUS CL974652
DEFINITION O51FC026296 Oryza sativa Express Library Oryza sativa (indica cultivar-group) genomic, genomic survey sequence.
ACCESSION CL974652
VERSION CL974652.1 GI:52403817
KEYWORDS GSS.
SOURCE Oryza sativa (indica cultivar-group)
ORGANISM Oryza sativa (indica cultivar-group)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzeae; Oryza.
REFERENCE 1 (bases 1 to 3201)
AUTHORS Ma,L., Wang,J., Chen,C., Liu,X., Su,N., Li,L., Wang,X., Cao,M., Jiao,Y., Sun,N., Zhang,X., Bao,J., Sun,D., Zhao,H., Yuan,L., Wong,G.K.S., Deng,X.W. and Wang,J.
TITLE An analysis of transcriptional regulation of the rice genome and its comparison to Arabidopsis
JOURNAL Unpublished (2004)
COMMENT Contact: Chen Chen
Department of Bioinformatic
Beijing Institute of Genomics
Chinese Academy of Sciences, Beijing 101300, China
Tel: 86-10-80481559
Fax: 86-10-80488676
Email: chenchen@genomics.org.cn
Rice genomic sequence.
Class: exon-trapped.
FEATURES
source
1..3201
/organism="Oryza sativa (indica cultivar-group)"
/mol_type="genomic DNA"
/db_xref="taxon:39946"
/clone_lib="Oryza sativa Express Library"
/note="Oryza sativa exon trapped genomic sequences"

```

## ORIGIN

```

Alignment Scores:
Pred. No.: 0.00163 Length: 3201
Score: 160.00 Matches: 80
Percent Similarity: 36.46% Conservative: 52
Best Local Similarity: 22.10% Mismatches: 128
Query Match: 4.42% Indels: 102
DB: 9 Gaps: 15

US-09-930-591-2 (1-686) x CL974652 (1-3201)
QY 183 ProValPheSerAspAsnSerSerProAlaValProGlnSerTyrGlnValAlaHis 202
Db 1252 CCAATATATAAATTCAGGATGAACCTGCTCAAGGCTGTTGAAGAGTATCAGGTATTGTC 1311
QY 203 LeuHisAlaProThrGlySerGlyLysSerThrLysValPro-----Ala 217
Db 1312 ATAGTAGGAGAACTCGCTCTGGTAAACACACAATAATACCTCAATATCTTCAAGAGCT 1371
QY 218 AlaTyAlaAlaGlnGlyTyLysValLeuValLeuAsnProSerValAlaAlaThrMet 237
Db 1372 GGNATACAGCAAAAGCA---AAGTGGCATGTACACAACCTCGTCGAGTTGCGCAATG 1428
QY 238 GlyPheGlyValaTyMetSerLys-----AlaHisGlyIleAspPro 251
Db 1429 AGTGTTCAGCGAGGGTGTCCCAAGAGATGGGTGTTAACTAGGACATGAGTTGGCTAC 1488
QY 252 AsnIleArg-----ThrGlyValArgThrIleThrThrGlySerProIleThr 267
Db 1489 TCTATAAGGTTTGAGGATTCGACGCTCTGAGAAAACATTG-----ATTAAAG 1533
QY 268 TyrSerThrGlyLysPheLeuAlaAsp-----GlyGlyCysSerGlyGlyAlaTy 285
Db 1534 TATACATCATGGCATGCTCTGAGGGAGTTCTTGGCGAACCAACATTTGGCAAGCTAT 1593
QY 286 AspIleIleCysAspGluCysHisSerThrAspAlaThrSerIleLeuGlyIleGly 305
Db 1594 AGTGTGTTATGGTTGATGAGGCTCATGAGCGTACACTGCTACTGATATCTTATTGGT 1653
QY 306 ThrValLeuAspGlnAlaGluThrAlaGlyAlaArgLeuThrValLeuAlaThrAlaThr 325
Db 1654 TTGGTGAAGGAT-----1665
QY 326 ProProGlySerValThrValProHisProAsnIleGluGluValAlaLeuSerThrThr 345
Db 1666 -----ATTTCAGGTTTCACAGATCTAAAGTTGCTCATTTCAAGTGCACACC 1713
QY 346 -----GlyGluIleProPheTyGlyLysAlaIleProLeuGluAlaIleLysGly 362
Db 1714 CTTGATGCAGAAAAATTTAGTGACTACTTTGATTGAGTCTCTATTTTCAAGATACCTGGG 1773
QY 363 GlyArgHisIleLeuPheCysHisSerLysLys-----LysCysAspGluLeuAlaAla 380
Db 1774 AGCGCATCTCTGTTGAAGTTTCATTATACAAAAGCTCCAGAGCAGATATATAGATGCA 1833
QY 381 LysLeuValAlaLeu-----385
Db 1834 GCCATTGTCACTGTTTTCAGATACATGTGACGCCAACCCCTGGTGATATCTTGTATTTC 1893
QY 386 -----GlyValAsnAlaValAlaTyTyTy-----ArgGly 395
Db 1894 CTTACAGGACNAGAAGAAATTGAACAATTGATGAATCCTTAAACACACAGACAGAGGC 1953
QY 396 LeuAspValSerVal-----400
Db 1954 TTAGGCACAAAGATTTCGAGAACTACTTATCTGTCCTATATATATGCAAAATCTGCCA 2013
QY 401 -----IleProThrSerGlyAspValValValValAlaThr 412
Db 2014 CTTCAAGCTAAGATATTGTAGCCACCCAGGAGGCTCGGAAGGTTGTTCTTGGCCACT 2073
QY 413 AspAlaLeuMetThrGlyPheThrGlyAsp---PheAspSerValIleAspCysAsnThr 431

```







300 rIleLeuGlyIleGlyThrValLeuAaspGlnAlaGluThrAlaGlyAlaArgLeuThrVa 320  
918 ACTTGTG---ACCGGTTTCAGTGGTTGTCTGGCCCTCTGGGTGGCATGGAGTGCACGT 974  
320 lLeuAla---ThrAlaThrPro----- 326  
975 GAGAAGTCAGACCGGATCCCCCAGATCCTCAAATGGCCCTCAGAATGGAGTTCAACTTA 1034  
327 -----ProGlySerValThr-----ValProHisPro----- 335  
1035 GAGACGATGCCCGGATCACTGTGCAGTCGACGGNACCCTTCCCGTGGCGGCAGC 1094  
336 -----AsnIleGluValAlaLeuSerThrThrGlyGluIleProPheTyrGlyLy 353  
1095 ATAGAGCTACGCAAGCCAGACGGCACTGTCTCTGTCCACCAAGGCCATTGTGGAGCCA 1154  
353 sAlaIlePro-----LeuGluAlaIleLysGlyGl 363  
1155 GAGAAGACCAAGCTGATGTCGAGGTGCCCGCTTGGTCTTCGGACAGATGGGTTCTGG 1214  
363 YArgHisLeuIlePheCysHisSerLysLysCysAspGluLeuAlaAlaLysLeuVa 383  
1215 GAGTGGCGTG-----GTGCCATCTGGCGCCAGACAGC 1250  
383 l-AlaLeu-GlyValAsnAlaValAlaTyrArgGlyLeuAaspValSerValIlePro 402  
1251 CGGCGCTTCAAGGTCAAT-----GTGAAAGTGCCCCC 1283  
403 ThrSerGlyAspValValValAlaThrAspAlaLeuMetThr----- 417  
1284 -----GTGCCCTTGGCTGCACTCGGCTCTGTGACCAAGCAGACGCCCCAG 1328  
418 -----GlyPheThrGlyAasp----- 422  
1329 CTTTGTGTCTCCCGTGGTCTCGTCTCTGGGGATGGACCCATCTCCACTGTCCGCCTG 1388  
423 -----PheAaspSerValIleAaspCysAsnThrCysValThrGlnThrValAasp 438  
1389 CACTACCGGCCCAGACGACATACCATGACTGGTCGACCAATTGTG----- 1433  
439 PheSerLeuAaspProThrPheThrIleGluThrIleThrIleu-----ProGln 454  
1434 -----GTGACCCCAAGT-----GAGAAGCTGACGTTAATGAACCTGAGGCCCAAAG 1478  
455 AspAlaValSerArgThrGlnArgAlaGlyArgThrGlyArgGlyLysProGlyIleTyr 474  
1479 ACAGGATACAGTGTTCGTGTGACGTAGCCGCCGCGAGGGGAAGGA----- 1523  
475 ArgPheValAlaProGlyGlu-ArgProSerGlyMetPheAaspSer-----SerVa 491  
1524 -----GGAGAGGGGGCTGGGGGCCCTCCACCCCTCATGACCACAGACTGT 1568  
491 lLeuCysGluCysTyrAspAlaGlyCysAla--TrrTyrGluLeuThrProAlaGluThr 510  
1569 CCTGAGCCTTTGT-----TGCACGCGTGTGTGGAGGGCTGGCATGTGGAAGGC 1616  
511 ThrValArgLeuArgAlaTyrMetAsnThrProGlyLeu-----ProValCysGlnAasp 528  
1617 ACTGACCGGCTGCAGTAGTACGTGTCCTTGGTCCCTTGGTCCCGGGCCACTGTGGGGCAG 1676  
529 His-----LeuGluPheTrpGluGly----- 535  
1677 GGTTCCTGTCGCTGTGGACGGGACACAGCGGGGCGAGAGCGGGGAGAACGCTCTCA 1736  
536 -----ValPheThrGlyLeuThrHisIleAspAlaHisPheLeu 548  
1737 TCCCCCAGCGCCGCACTGCCCTCTCTGACGGGACTCACGCCT---GGCACCCACTACCAAG 1793  
549 SerGlnThrLysGlnSerGlyGluAsnLeuProTyrLeuValAlaTyrGlnAlaThrVal 568  
1794 CTGAGTGTGAG-----CTCTACCACCTGCACCCCTC 1823  
569 CysAlaArgAlaGlnAlaPro-----ProProSerTrpAaspGlnMetTrp 583

```

Db      1824 CTGGCCCGCCCTCGCCCTCGACACGTGCTTCTGCCCCCGAGTGGGCT----- 1874
Qy      584 LysCysLeuIleArgLeuLysProThrLeuHisGlyProThrProLeu---LeuTyArg 602
Db      1875 -----CCAGCCCCCGACACCTCCAGCC 1898
Qy      603 LeuGlyAlaValGlnAenGluValThrLeuThr-----His 614
Db      1899 CAGGCCCTTCACACTCCGAGATCCAGCTGACATGGAAGCACCAGCGGCTCTGCCCTGGG 1958
Qy      615 ProValThrLysTyrlleMetThrCysMetSerAla----- 626
Db      1959 CCAATATCAAGTACGTGTGGAGGTGCAGGTGCTGGGGTGCAGGAGACCCACTGTGG 2018
Qy      627 ---AspLeuGluValValThrSerThrTrpValLeuValGlyGlyValLeuAla 644
Db      2019 ATAGACGTGACAGCGCTGAGGAGACAGCACCATCATCGTGGCTCAACGCCAGC 2075

RESULT 6
LOCUS   BU054791
DEFINITION   822 bp mRNA linear EST 26-AUG-2002
IMAGE:6404900 5', mRNA sequence.
ACCESSION   BU054791
VERSION     BU054791.1 GI:22494868
KEYWORDS    EST.
SOURCE      Mus musculus (house mouse)
ORGANISM    Mus musculus
REFERENCE   1 (bases 1 to 822)
AUTHORS    NIH-MGC http://mgi.nhl.nih.gov/
TITLE      National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL    Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgabs-r@mail.nih.gov
            Tissue Procurement: Dr. James Lin, University of Iowa
            CDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
            DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
            Clone Distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            This clone was contributed by the Brain Molecular Anatomy Project
            (BMAP)

Seq primer: pyX-5.
Location/Qualifiers
1. .822
   /organism="Mus musculus"
   /mol_type="mRNA"
   /strain="C57BL/6"
   /db_xref="taxon:10090"
   /clone="IMAGE:6404900"
   /tissue_type="whole brain"
   /dev_stage="embryo 12.5 dpc"
   /lab_host="DH10B (T1 phage resistant)"
   /clone_lib="NIH-EMAP_FDO"
   /notes="Organ: brain; Vector: pyX-Asc; Site 1: EcoR I;
Site 2: Not I; The library was constructed according to
Bonaldi, Lennon and Soares, Genome Research, 6:791-806,
1996. Denatured mRNA was size fractionated on a 1% agarose
gel. First strand cDNA synthesis was primed with an
oligo-dT primer containing a Not I site. Double stranded
cDNA was size selected according to mRNA size fraction,
ligated with EcoR I adaptor, digested with Not I, and then
cloned directionally into pyX-Asc vector. The library tag
sequence located between the Not I site and the polyA
tail, is TGACAGAGCC. This library was created for the
University of Iowa Mouse Brain Molecular Anatomy Project
(EMAP): 'Gene Discovery in the Developing Mouse Nervous
System', supported by National Institutes of Mental Health
(NIMH), Hemin Chin, Ph.D., program coordinator."

```

## ORIGIN

```

Alignment Scores:
Pred. No.: 0.0146 Length: 822
Score: 140.50 Matches: 73
Percent Similarity: 35.33% Conservative: 45
Best Local Similarity: 21.86% Mismatches: 129
Query Match: 3.88% Indels: 87
DB: 5 Gaps: 11

US-09-930-591-2 (1-686) x BU054791 (1-822)
Qy      98 CysThrCysGlySerSerAspLeuValThrArgHisAlaAspValIleProVal 117
Db      12 TCGTGTGCGCAGGAGGCAAGTACCAGCTGGTGTGGAGGAGCAGACCATCGAGTTT 71
Qy      118 ArgArgArgGlyAspGlyArgGlySerLeuLeuSerProArgProIleSerTyrlleLys 137
Db      72 GTCCGTGCTGCTCAACTCCAGGGTGAC-----GAG 101
Qy      138 GlySerSerGlyGlyProLeuLeuCysProAlaGlyHisAlaValGlyIlePheArgAla 157
Db      102 GAGCCGTAGGCCCGCCCTGTGAGCCAGGCCAGCAGAAG----- 143
Qy      158 AlaValCysThrArgGlyValAlaLysAlaValAspPheIleProValGluSerLeuGlu 177
Db      144 -----GAATCTATCCAG 155
Qy      178 ThrThrMetArgSer---ProValPheSerAspAsnSerSerProProIleValProGln 196
Db      156 GCTGTGCGCGCAGCGCTGCGCTGTTCCTCCGAGAGAGCTTCTGCCCGCCATGTGCC 215
Qy      197 SerTyrlleValAlaHisLeuHisAlaProThrGlySerGlyLysSerThrLysValPro 216
Db      216 AACCATCAGTCTCTATCATCGAAGCGGAGACTGGCTCTGGGAGACACACAGATCCCA 275
Qy      217 -----AlaAlaTyrlleAlaGlnGlyTyrlleLysValLeuValLeuAsnPro 231
Db      276 CAGTACCTCTTTGAGGAGGGTTTACACAAAGAGGGCATGAAGATTGCTTGCACCCAGCCC 335
Qy      232 SerValAlaAlaThrMetGlyPheGlyAlaTyrlleMetSerLysAlaHisGly----- 248
Db      336 CGGAGAGTGGCGGCTATGAGTGTGGCAGCCCGAGTGGCCCGGAGAGATGGTGTGAAGCTT 395
Qy      249 -----IleAspProAsnIleArg-----ThrGlyValArgThrIleThr 261
Db      396 GGAACGAGGTGGGTACAGCATCCGGTTTGAGGACTGCACCTCAGAGAGTCTCTCTGAG 452
Qy      262 ThrGlySerProIleThrTyrlleThrTyrlleGly-----LysPheLeuAlaAsp 277
Db      453 -----CTCCGCTACATGACAGATGGAATGCTACTCTCCGAGAGTCTCTCTCTGAG 500
Qy      278 GlyGlyCysSerGlyAlaTyrlleAspIleIleCysAspGluCysHisSerThrAsp 297
Db      501 CCTGACCTTGCA-----AGTTACAGTGTGGTGAATGATGAAGACTCAGAGCGGAGCC 554
Qy      298 AlaThrSerIleLeuGlyIleGlyThrValLeuAspGlnAlaGluThrAlaGlyAlaArg 317
Db      555 TTGCACACACACATCTCTTTGATGATGAAGAGCTCGTAGATCCGACCTCGAGCTC 614
Qy      318 LeuThrValLeuAlaThrAlaThrProGlySerValThrValProHisProAsnIle 337
Db      615 AAGGTCTGTGGTTCAGCCACA----- 638
Qy      338 GluGluValAlaLeuSerThrThrGlyGluIleProPheTyrlleGlyLysAlaIleProLeu 357
Db      639 -----CTGATACTGCCCGTTCCTGCTTCTTCGATGACCCCTGTCTTCA 686
Qy      358 GluAlaIleLysGlyLysArgHisLeu---IlePheCysHisSerLysLysCysAspG 377
Db      687 GAATCTTGACCGCAGCTTCCAGTTGACATCTTATACCAAGGCCCGCCAGAGCTGACT 746
Qy      377 IuLeuAlaLysLeuValAlaLeuGlyValAsnAlaValAlaTyrlleArgGlyLeuA 397

```

```

Db      747 ACCCTGGAGCTGGCTGGTCTGG-----CTGC 776
Qy      397 spValSerValIleProThrSerGlyAspValValVal 409
Db      777 AGATTACGTCAGCCAGCCCTGGAGATACACTGGTG 814

RESULT 7
BC059369
LOCUS
DEFINITION Homo sapiens cDNA clone IMAGE:30346256, containing frame-shift
errors.
ACCESSION BC059369
VERSION BC059369.1 GI:37590735
KEYWORDS HTC.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 3320)
Strauberg,R.L., Feingold,B.A., Grouse,L.H., Derge,J.G.,
Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,
Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,
Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,
Distchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,
Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S.,
Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,
Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J.,
McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,
Villalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
Fahey,J., Helton,E., Kettelman,M., Madan,A., Rodrigues,S.,
Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y.,
Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,
Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
Butterfield,Y.S., Krzywicki,M.I., Skalska,U., Smallos,D.E.,
Schnerch,A., Schein,J.B., Jones,S.J. and Marra,M.A.
Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
22388257
1247932
2 (bases 1 to 3320)
Strauberg,R.
Direct Submission
Submitted (01-OCT-2003) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NIH-MGC Project URL: http://mgc.nci.nih.gov
Contact: MGC help desk
Email: cgapbs-remail.nih.gov
Tissue Procurement: Dr. Stefan Hansson
cDNA Library Preparation: Michael Brownstein / Ted Usdin
Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILLNL)
DNA Sequencing by: Sequencing Group at the Stanford Human Genome
Center, Stanford University School of Medicine, Stanford, CA 94305
Web site: http://www-shgc.stanford.edu
Contact: (Dickson, Mark) mcd@paxil.stanford.edu
Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers,
R. M.

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/ILLNL at: http://image.llnl.gov
Series: IRAC Plate: 133 Row: i Column: 8
This clone was selected for full length sequencing because it
passed the following selection criteria: matched mRNA gi: 20544128
This clone has the following problem: frame shifted.
Location/Qualifiers
1. .3320
/organism="Homo sapiens"

FEATURES
source

```

```

/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:30346256"
/tissue_type="placenta, normal"
/clone_id="NIH MGC_147"
/lab_host="DH10B"
/note="Vector: pbluescript"

```

## ORIGIN

```

Alignment Scores:
Pred. No.: 0.306 Length: 3320
Score: 136.50 Matches: 114
Percent Similarity: 33.69% Conservative: 73
Best Local Similarity: 20.54% Mismatches: 216
Query Match: 3.77% Indels: 153
DB: 3 Gaps: 24

US-09-930-591-2 (1-686) x BC059369 (1-3320)
Qy 132 ProfileSerTyrLeuLysGlySerSerGlyProLeuLeuCysProAlaGlyHisAla 151
Db 35 CCGGTGAAGTTCTGGCGACCCGGTACAGAGGGCCA-----GGTGTAAAGC 79
Qy 152 ValGlyIlePheArgAlaAlaValCysThrArgGlyValAlaLysAlaValAspPheIle 171
Db 80 ATCTCTGAAGAGAGACAAAGTCTGGCTGAAACTCTCT---GGGACAAACGGTTGTTTACAAC 136
Qy 172 ProValGluSerLeu-----GluThrThrMetArgSerProValPheSerAspAsn 188
Db 137 CCTTATGCTGCCCTTTCATAGACAGCAGAGGAGAGTGGCGGTATTCAAGCTTAGG 196
Qy 189 SerSerProProAlaValProGlnSerTyrGlnValAlaHisLeuHisAlaProThrGly 208
Db 197 AATCATATTTTATCTTGTAGATAAATTAACAGCGGTGGTGTGTTGGTGAACAGGA 256
Qy 209 SerGlyLysSerThrLysValPro-----AlaAlaTyrAlaAlaGlnGly 223
Db 257 TGTGGGAAGAGCAGACAGATTCTCAGTACCTTCGAGAAGCCGGCTGCAGCAGCTGAAGGA 316
Qy 224 TyrLysValLeuValLeuAsnProSerValAlaAlaThrMetGlyPheGlyAlaTyrMet 243
Db 317 AGAGTGTGAGGAGTGACCCAGCCTCGAAGAGTGCTGCTGTACAGTTGTCAGGAGAGTA 376
Qy 244 SerLysAlaHisGlyIleAspProAsnIleArgThrGlyValArgThrIleThrThrGly 263
Db 377 GCTGAACAAGGGGTGC-AGTGTGGGCCACAGAGTGGGCTACTG-----420
Qy 264 SerProIle-----ThrTyrSerThrTyrGlyLysPheLeuAlaAspGly 278
Db 421 CATCCGCTTTTGATGACTGCACCGACGACGTGCCACGAGAATTAAGTTTCTTACTGTGA 480
Qy 279 GlyCysSerGly-----GlyAlaTyrAspIleIleIleCys 290
Db 481 ATGCTGCTCAGGGAATATGATGGTTGATCCGTTGTTAAACAAATATAGTGTATCATGTG 540
Qy 291 AspGluCysHisSerThrAspAlaThrSerIleLeuGlyIleGlyThrValLeuAspGln 310
Db 541 GATGAAGCCACGAGAGGACCTTGTCACATGCCATTGGCTTGGCTTGAATAAAGATT 600
Qy 311 AlaGluThrAlaGlyAlaArgLeuThrValAlaThrAlaThrProGlySerVal 330
Db 601 CAGAAAAAGCGAGGGGATCTTCGATTGATTGATTCAGCCCTCTCGGATGCAGACAAA 660
Qy 331 ThrValProHisProAsnIleGluGlu-----ValAla 341
Db 661 TTCGGGATTTCTTTAATCAAAATGAACCAAGTATCCAGCAAGGGATACATGTGTGATC 720
Qy 342 LeuSerThrThrGlyGlu-----IlePro 349
Db 721 CTTACAGTGAAGGAGAACATTTCCGGTGGATATCTTTTATCTACAAAGTCTCTGTCCA 780
Qy 350 PheTyrGlyLysAlaIleProLeuGluAlaIleLys-----GlyGly 363

```

Db 781 GATTATATCAAAATCAACTGTCGAAACTGTGGTGAATAATTCACAGACAGAGGGAGACGA 840  
 Qy 364 ArgHisLeuIlePheCysHisSerLysLysLysCysAspGluLeuAlaLysLeuVal 383  
 Db 841 GACGTTTACCAATTTCTACTGCCAGGAGAGGTGAAACTGTTGTGTCGATGCTCATC 900  
 Qy 384 -----AlaLeuGly-----ValAsnAlaValAlaLys 392  
 Db 901 GACCAGGCTCGACACTAGCTCGCACTGGGATGAGAGACACCTCCGAGTCTCCCCATG 960  
 Qy 393 TyrArgGlyLeuAsp-----ValSerValIleProThrSerGly 405  
 Db 961 TATCGAGGACTGCTCTTCTTGGCAAAATGAAAGTGTGTTGAAAGGTGTGCACGACTGC 1020  
 Qy 406 AspValValValAlaLysAlaLeuMetThrGlyPheThr---GlyAspPheAsp 424  
 Db 1021 AGAAAGTGATAGTGGCCCAATGTCGCAAAATGTCGCAAACTCTATCAATCAGCGGCATTGTG 1080  
 Qy 425 SerValIleAspCysAnthrCysValThrGlnThrValAspPheSerLeuAspProThr 444  
 Db 1081 TATGTGATCGACTGTGCTTGTGMAACTCCGA-----GCCATCAATCCAGG 1128  
 Qy 445 PheThrIleGlu---ThrIleThrLeuProGlnAspAlaValSerArgThrGlnArgArg 463  
 Db 1129 ACAGCTATTGAATGCTTGTGTGTGTCGCCAGTCTCCCAAGCATCAGCTAATCAGCGAGCA 1188  
 Qy 464 GlyArgThrGlyArgGlyLysProGly---IleTyrArgPheValAlaProGlyGluArg 482  
 Db 1189 GGACGTGTGTGCTAGTCTCGGAAATGTTATGCTCTTATACA-----GAGGAA 1242  
 Qy 483 ProSerGlyMetPheAspSerValLeuCysGlyCysTyrAspAlaGlyCysAlaLys 502  
 Db 1243 GCCTTTGACAGTGTGCTCAGTCTACGTTCTGAGATGAGCGTAGTAATTG----- 1296  
 Qy 503 TyrGluLeuThrProAlaGluThrThrValArgLeuArgAlaTyrMetAsnThrProGly 522  
 Db 1297 -----GCACCTGTCTATCTGTCAGCTGAAAGCA----- 1323  
 Qy 523 LeuProValCysGlnAspHisLeuGluPheThrGlyValPheThrGlyLeuThrHis 542  
 Db 1324 -----CTAGGAATGTCAAT 1338  
 Qy 543 Ile---AspAlaHisPheLeuSerGlnThrLysGlnSerGlyGluAsnLeuProTyrLeu 561  
 Db 1339 GTCTCAGGTTCACCTTCATGTCG----- 1362  
 Qy 562 ValAlaTyrGlnAlaThrValCysAlaArgAlaGlnAlaProProSerThrAspGln 581  
 Db 1363 -----CCCCCTCCAGCA---CAGTCG 1380  
 Qy 582 MetTrpLysCysLeuIleArgLeuLysProThrLeuHisGlyProThrProLeuLeuTyr 601  
 Db 1381 ATGGTTCAAGCCTTG-----GAGTTACTGTAT 1407  
 Qy 602 ArgLeuGlyAlaValGlnAsnGluValThrLeuThrHisProVal 616  
 Db 1408 GCTCTGGGAGGTCTGGCAAGACTGTGCGCTAACTGAACCGCTT 1452

RESULT 8  
 CD359697  
 LOCUS  
 DEFINITION AGENCOURT 14284682 NIH MGC 180 Homo sapiens cDNA clone  
 IMAGE:30388971 5', mRNA sequence.  
 ACCESSION CD359697  
 VERSION CD359697.1 GI:31131108  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 REFERENCE 1 (bases 1 to 889)  
 AUTHORS NIH-MGC <http://mgi.nci.nih.gov/>.  
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL  
COMMENT

Unpublished (1999)  
 Contact: Daniela S. Gerhard, Ph.D.  
 Office of Cancer Genomics  
 National Cancer Institute / NIH  
 Bldg. 31 Rm10A07 Bethesda, MD 20892  
 Email: cgebs-r@mail.nih.gov  
 Tissue Procurement: Dr. Michael Brownstein  
 cDNA Library Preparation: Invitrogen Corp  
 DNA Sequencing by: The I.M.A.G.E. Consortium (LNL)  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LNL at:  
<http://image.llnl.gov>  
 Plate: NDAM461 row: 1 column: 04  
 High quality sequence start: 22  
 High quality sequence stop: 653.

FEATURES  
source

1. .889  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:30388971"  
 /lab\_host="DH10B-Ton A (T1 and T5 phage resistances)"  
 /clone\_lib="NIH\_MGC\_180"  
 /note="Organ: Testis; Vector: pCMV-SPORT6.1; Site 1: NotI;  
 Site 2: EcoRV (destroyed); Library is oligo-dT primed and  
 directionally cloned (EcoRV site is destroyed upon  
 cloning). Average insert size 1.68 kb. Library was  
 constructed by (Invitrogen). Note: this is a NIH\_MGC  
 Library."

## ORIGIN

Alignment Scores:  
 Pred. No.: 0.0771 Length: 889  
 Score: 133.50 Matches: 61  
 Percent Similarity: 40.24% Conservative: 40  
 Best Local Similarity: 24.30% Mismatches: 95  
 Query Match: 3.69% Indels: 55  
 DB: 6 Gaps: 11

US-09-930-591-2 (1-686) x CD359697 (1-889)

Qy 183 ProValPheSerAspAsnSerSerProAlaValProGlnSerTyrGlnValAlaHis 202  
 Db 44 CCGGTGTTCCCATTTCCAGAGGAGCTCTGGCTGCTATTGCAATCACCAGTCTCATC 103  
 Qy 203 LeuHisAlaProThrGlySerGlyLysSerThrLysValPro-----Ala 217  
 Db 104 ATTGAAGGCGGAGACAGCTCAGGGAAGACCCAGATCCCGCAGTATCTCTTTGAGGAG 163  
 Qy 218 AlaTyrAlaAlaGlnGlyTyrLysValLeuValLeuAsnProSerValAlaAlaThrMet 237  
 Db 164 GGTATATCAAAACAGGATGATGAGATGCTGCACCAACCCCGGAGAGTGGTGCATG 223  
 Qy 238 GlyPheGlyAlaTyrMetSerLysAlaHisGly-----IleAspPro 251  
 Db 224 AGTGTGCGCCCGAGTGGCCCGGAGATGGGTGTGAAGCTTGGGAATGAGTTGGCTAC 283  
 Qy 252 AsnIleArg-----ThrGlyValArgThrIleThrThrGlySerProIleThr 267  
 Db 284 AGCATCCGCTTTGAGGACTGCATCATCAGACGCAACTGTC-----CTCCGC 328  
 Qy 268 TyrSerThrTyrGly-----LysPheLeuAlaAspGlyGlyCysSerGlyGly 283  
 Db 329 TACATCAGATGGGATGCTTCTCCGGAGTCTCTCTGAGCTGACCTGGCG----- 382  
 Qy 284 AlaTyrAspIleIleCysAspGluCysHisSerThrAspAlaThrSerIleLeuGly 303  
 Db 383 AGTTACAGCTGTGTGTGTGGATGAGCACACAAAGGACCCCTACACAGACATCTTC 442  
 Qy 304 IleGlyThrValLeuAspGlnAlaGluThrAlaGlyAlaArgLeuThrValLeuAlaThr 323  
 Db 443 TTTGGATGTATCAAGGATGTTGCTCGCTCCGACCTCAGCTCAAGGTCTCTGGTGGCTTCA 502









Qy 653 GlyCysValValIleValGlyArgIleValLeuSerGlyLysProAlaIleIleProAsp 672  
 Db 2368 GGCTGGCTCCCTCCGAGAGCGGCTGTCAGGCGCACTCCCTAGTGCCTCC 2427  
 Qy 673 Arg 673  
 Db 2428 AGG 2430

RESULT 12  
 AK031534  
 LOCUS  
 DEFINITION Mus musculus 13 days embryo male testis cDNA, RIKEN full-length enriched library, clone:6030448F20 product:hypothetical P-loop containing nucleotide triphosphate hydrolases structure containing protein, full insert sequence.

ACCESSION AK031534  
 VERSION AK031534.1 GI:26327388  
 KEYWORDS HTC; CAP trapper.  
 SOURCE Mus musculus (house mouse)  
 ORGANISM Mus musculus

REFERENCE 1  
 Carninci, P. and Hayashizaki, Y.  
 TITLE High-efficiency full-length cDNA cloning  
 JOURNAL Meth. Enzymol. 303, 19-44 (1999)  
 MEDLINE 92279253  
 PUBMED 10349636

REFERENCE 2  
 Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.  
 TITLE Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes  
 JOURNAL Genome Res. 10 (10), 1617-1630 (2000)  
 MEDLINE 20499374  
 PUBMED 11042159

REFERENCE 3  
 Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Konno, H., Akiyama, J., Nishi, K., Kiseunai, T., Taehiro, H., Itoh, M., Suni, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwaka, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsui, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.  
 TITLE RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer  
 JOURNAL Genome Res. 10 (11), 1757-1771 (2000)  
 MEDLINE 20530913  
 PUBMED 11076861

REFERENCE 4  
 The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium.  
 TITLE Functional annotation of a full-length mouse cDNA collection  
 JOURNAL Nature 409, 685-690 (2001)

REFERENCE 5  
 The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.  
 TITLE Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs  
 JOURNAL Nature 420, 563-573 (2002)

REFERENCE 6  
 (bases 1 to 2972)  
 Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T., Kato, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N., Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A., Muramatsu, M. and Hayashizaki, Y.

TITLE JOURNAL  
 COMMENT  
 FEATURES  
 source  
 1. Location/Qualifiers  
 /organism="Mus musculus"  
 /mol\_type="mRNA"  
 /strain="C57BL/6J"  
 /db\_xref="PANTOM\_DB:6030448F20"  
 /db\_xref="taxon:10090"  
 /clone="6030448F20"  
 /sex="male"  
 /tissue\_type="testis"  
 /clone\_lib="RIKEN full-length enriched mouse cDNA library"  
 /dev\_stage="13 days embryo"  
 17. .2641  
 /note="unnamed protein product; hypothetical P-loop containing nucleotide triphosphate hydrolases structure containing protein (SCOP|52540, evidence: SCOP) putative"  
 /codon\_start=1  
 /protein\_id="BAC27438.1"  
 /db\_xref="GI:26327389"  
 /translation="MAAGRRVGPPRPTMAPKKKRLKRRRTGASQGRSDSDGGEFVABDDARAKLGPGRALPGPPTSECVSDVEPDTRVMVRAQNKKKKSGGQSMGLSYEVFKGIMKGVKVPTRQRTIPVLGDKDVAMARTSGTKACFLFLPMPRLKARSAQTGARALLSPTRELALQTKFKELGKFTGLKTLALILGDKMEDQFAALHENDPIIATPGRVHVAVMNKLQSVYVDEADRLFEWGFAGLQRIIGRPGHQTVDLSATPLKLVFAQAGLEPVLRLVDKLNELQKTSFLLRVEDTAAVLLYLQNVVRQDQTVFVATKHAELYELTMGGQSCAHYISALDQTKARKINAKFTHNKCSSTLI VTDLAARGDIPLDNVIIVSPAKGLFLHRVGRVARAGSGTASLVADPVPVILLDIHLFGRSVTIARPCPEPSADAVGRDGVLRVQSVVDDEDSLTQAMGASLDLQGLHRVANNAAQQVVRSPAPSPSILKAKELDLAELGLPLFSSCFEGELQRLVDSIKNRTITTIIFELNASSKDPSSQMMRAKQRKAVASQQRRQERQEGPADPAPQRELPOEEEEENVEVTEGVFTVVGQRPRPGSGAKRRMETRORDQEFYVYPRKPDFSERGLSVGAGGAFEQQVAGVLDLMDGAEQNMRSQQQLKWDKRRKRVFGSGQEDKKITKTSGRFISSSYKRDLYQWKQKIDRDSDEEGPSNQRPGPRRGRGGRSQGTSPRASSVPAGMRSELTKTEQILQRRQAQKQRFQRLQGLKQLSARNRRRAQLRQGAFGRAPSRKGNKEM"

polyA\_signal  
 2954. .2959  
 /note="putative"  
 2972  
 /note="putative"

ORIGIN  
 Alignment Scores:  
 Pred. No.: 1.34 Length: 2972  
 Score: 129.00 Matches: 134  
 Percent Similarity: 31.24% Conservative: 55  
 Best Local Similarity: 22.15% Mismatches: 216  
 Query Match: 3.57% Indels: 201  
 DB: 3 Gaps: 29

US-09-930-591-2 (1-686) x AK031534 (1-2972)  
 Qy 100 CysGlySerSerAspLeuTyrLeu-----ValThrArgHisAlaAspVal 114  
 Db 90 TCGGAAACCCCACTGGGGCTTCCCAAGCCGCGACACCGACTCGATGACGGCAGT 149  
 Qy 115 IleProValArgArg-----GlyAspGlyArgGlySerLeuLeuSerProArg 131

```

Db      150 TCAGATTCAGCGGAGGATGACGCCGGCGGAGGAAGCTGGCCCTGGCAGAGCTTGC 209
Qy      132 Pro-IleSerTyrLeuLys-----GlySerSerGlyGlyProLeuLeu 145
Db      210 CCTCATTTCTTACTCAGAGTGGGTATCAGATGTGGAGCCGCACACTCGGAGATGGTGC 269
Qy      145 uCysProAlaGlyHisAlaValGlyIlePheArgAlaAlaVal----- 159
Db      270 GAGCCACAGAACAGAAAAAGAAAGAGTCTGGAGGCTTCAGTCCATGGGCTCAGTTACC 329
Qy      160 -CysThr-ArgGlyValAlaLysAlaValAspPheIleProValGluSerLeuGluThrT 179
Db      330 CTGTTTCAAGGGGATCATGAAAAAGGGCTACAAGGTGCCGACGCCCATCCAGAGAAAGA 389
Qy      179 hrMetArgSerProValPheSerAspAsnSerSerProProAlaValProGlnSerTyrG 199
Db      390 CCATC-----CCCTGTATCTTGGATGGCAAGGATGGTGCCATG----- 430
Qy      199 InValAlaHisLeuHisAlaProThrGlySerGlyLysSerThr----- 213
Db      431 -----GCCCGACAGCGAGTGGCAAGACGCGCTGTCTCTCTCCCGA 473
Qy      214 -----LysValProAlaAlaTyrAlaAlaGlnGlyTyrLysValLeuValLeuAsnP 231
Db      474 TGTTTGAGCGGTGAAGGACGACGTGCACAGACGGGGGCTCGAGCCCTCATCTCTCTCAC 533
Qy      231 roSer-----ValAlaAlaThrMetGlyPheGlyAlaTyrMetSerLysAlaHisG 248
Db      534 CCACCGGGAGCTGCCCTCGACACCATGAGTTCACTTAAGAGTAGGCAAGTT-CACC 592
Qy      248 ly-----IleAspProAsnIleArgThrGlyValArgThrIleThrThrGlys 264
Db      593 GGCCTCAAGACTGCCTTGCATCTGGTGGAGACAAATGGAAGACCACTTTGCAGCCCTG 652
Qy      264 erProIleThrTyrSerThrTyrGlyLysPheLeuAlaAspGlyGlyCysSerGlyGlyA 284
Db      653 CACGAGAACCTTGACATAATCATTCACCCCTCGGGCGTCTGGTGCATGT-----GGCT 706
Qy      284 laTyrAspIleIleCysAspGluCysHisSerThrAspAlaThrSerIleLeuGlyI 304
Db      707 GTGGAGTGAATGAACTGAAGCTGCAGAGTGT-----GGAGTATGTGGTG 748
Qy      304 le----- 304
Db      749 TTGATGAAGCAGACAGGCTCTTTGAATGGCTTTGCTGAGCAGCTACAGGATCATATA 808
Qy      305 -----GlyThrValLeuAsp---GlnAlaGluThrAlaG 315
Db      809 GCGCGCTTCTTGGGGCCACAGACGCTGTCTTCTCAGCTACACTGCCCAAGCTGCTG 868
Qy      315 lyAlaArgLeuThr-----ValLeuAlaThrAlaThrProProGlySerValThrV 332
Db      869 GTGGAAATTCACAGCGAGGCTCACAGAGCCCGTCTCATCCGCTCGAGCTAGACTCC 928
Qy      332 alProHisProAsnIleGluGluValAlaLeuSerThrThrGlyGluIlePro----- 349
Db      929 AAGCTCAATGACAGCTCAAGACCTC-----CTTCTCTCTTGTG 967
Qy      350 -----PheTyrGlyLysAlaIleProLeuGluAlaIleLys-----GlyG 363
Db      968 CGCGAAGACACCAAGCTGTGTGCTCTCTCTACCTGTGTCAGATGTGTTGGGCCCCAG 1027
Qy      363 ly-ArgHisLeuIlePheCysHisSerLysLysLysCysAspGluLeuAlaAlaLysLeu 382
Db      1028 GACCAGACTGGGTGTGTAGCCACCAAGACCAATCGCGAGTAGTCTCACAGATGTGCTG 1087
Qy      383 ValAlaLeuGlyValAsnAlaValAlaTyrTyrArgGlyLeuAspValSerValIlePro 402
Db      1088 ATGGCCAGGTGTGAGTGTGGCCCACTATATAGTGTGCTTGGACAGAC----- 1138
Qy      403 ThrSerGlyAspValValValAlaThrAspAlaLeuMetThrGlyPheThrGlyAsp 422

```

```

Db      1139 -----GCCCGCAAGATCAACTTGGCCCAAGTTTCACACACAAC 1174
Qy      423 PheAspSerValIleAspCysAsnThr---CysValThrGlnThrValAspPheSerLeu 441
Db      1175 -----AAATGTTCCACCTCATCTGACTGACCTGGCCGCCGGGGCGCTG 1219
Qy      442 AspProThrPheThrIleGluThrIleThrLeuProGlnAspAla----- 456
Db      1220 GACATCCCACTGCTGGCAACGCTCATCACTACAGCTTCCCTGCCAAGGGCAAGCTCTTC 1279
Qy      457 -----ValSerArgThrGlnArgArgGlyArgThrGlyArgGlyLysProGlyIle 473
Db      1280 CTGACCCAGTGGCGCGTGTGGCCGACAGCGCCGCAAGTGGCACAGCC----- 1327
Qy      474 TyrArgPheValAlaProGlyGlu-ArgProSerSerGlyMetPheAspSerSerValLeuCy 493
Db      1328 TATCTTTGGTGGCCCCAGACGAGGTCCCTTACC----- 1361
Qy      493 sGluCysTyrAspAlaGlyCysAlaTyrTyrGluLeuThrProAlaGluThrThrValAr 513
Db      1362 ---TGCTTGACCTACACCTGTCTGGCGCGCTGTGTACCTGGCCGCTCTTGTGAG 1417
Qy      513 g---LeuArgAlaTyrMetAsnThrProGlyLeuProValCysGlnAspHisLeuGluPh 532
Db      1418 GAGCTTCAGTGGCAGATGCGTGGCAGGG-----ACGAGGTG 1456
Qy      532 eTyrGluGly----- 535
Db      1457 CTGGTCCGTCGCCAGAGTGTAGTGATGATGAGGACAGCAGCTCGAGCTCCCATG 1516
Qy      536 -----ValPheThrGlyLeuThrHisIleAspAlaHisPheLeuSerGlnTh 551
Db      1517 GGGGCATCCCTGGATCTTCAGGGCGTCGACC----- 1547
Qy      551 rLysGlnSerGlyGluAsnLeuProTyrLeuValAlaTyrGlnAlaThrValCysAlaAr 571
Db      1548 -----GGTGGCCCAACAACGCTCAGCAGCAGTATGTGGCTCA 1585
Qy      571 gAlaGlnAlaPro-----ProProSerTyrAsp---GlnMetTr 583
Db      1586 CGGCAGCGCTCGCTGAGTCCATCAAGAGAGCCAGGAGCTGGACCTGGCAGAGCTG 1645
Qy      583 pLysCys 585
Db      1646 GGCTTGC 1652

RESULT 13
AK031679
LOCUS
DEFINITION
Mus musculus 13 days embryo male testis cDNA, RIKEN full-length
enriched library, clone:6030486C03 product:hypothetical P-loop
containing nucleotide triphosphate hydrolases structure containing
protein, full insert sequence.
ACCESSION
AK031679
VERSION
AK031679.1 GI:26327530
KEYWORDS
HTC; CAP trapper.
SOURCE
Mus musculus (house mouse)
ORGANISM
Mus musculus

REFERENCE
1
Carninci, P. and Hayashizaki, Y.
High-efficiency full-length cDNA cloning
Meth. Enzymol. 303, 19-44 (1999)
99279253
PUBMED
10349636
2
Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
Itoh, M., Komo, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes
Genome Res. 10 (10), 1617-1630 (2000)
20499374
PUBMED

```

```

11042159
3
PUBMED
REFERENCE
AUTHORS
Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,
Konno, H., Akiyama, J., Nishi, K., Kiseunai, T., Tashiro, H., Itoh, M.,
Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Harada, A.,
Fujiwaka, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M.,
Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuura, S., Kawai, J.,
Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
RIKEN integrated sequence analysis (RISA) system-394-format
sequencing pipeline with 394 multicapillary sequencer
Genome Res. 10 (11), 1757-1771 (2000)
20530913
11076861
4
The RIKEN Genome Exploration Research Group Phase II Team and the
FANTOM Consortium.
Functional annotation of a full-length mouse cDNA collection
Nature 409, 685-690 (2001)
5
The FANTOM Consortium and the RIKEN Genome Exploration Research
Group Phase I & II Team.
Analysis of the mouse transcriptome based on functional annotation
of 60,770 full-length cDNAs
Nature 420, 563-573 (2002)
6 (bases 1 to 2984)
Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P.,
Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W.,
Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T.,
Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T.,
Kato, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M.,
Koya, S., Kurihara, C., Matsuura, T., Miyazaki, A., Murata, M.,
Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N.,
Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N.,
Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T.,
Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akai, S.,
Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A.,
Muramatsu, M. and Hayashizaki, Y.
Direct Submission
Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of
Physical and Chemical Research (RIKEN), Laboratory for Genome
Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
Kanagawa 230-0045, Japan (E-mail: genome-res@gsc.riken.jp,
URL: http://genome.gsc.riken.jp/, Tel: 81-45-503-9222,
Fax: 81-45-503-3216)
cDNA library was prepared and sequenced in Mouse Genome
Encyclopedia Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in RIKEN.
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues
Please visit our web site for further details.
URL: http://genome.gsc.riken.jp/
URL: http://fantom.gsc.riken.jp/.
FEATURES
source
1. .2984
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db xref="FANTOM DB:6030486C03"
/db xref="taxon:10090"
/clone="6030486C03"
/sex="male"
/tissue type="testis"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev stage="13 days embryo"
25. .2754
/notes="unnamed protein product; hypothetical p-loop
containing nucleotide triphosphate hydrolases structure
containing protein (SCOP|52540, evidence: SCOP)
putative"
/codon_start=1
/protein_id="BAC27509.1"

```

```

/db_xref="GI:26327531"
/translation="MAAGRRVGPPSPRPTMAPWKKRLKRLKRTGTASQGRDSDSDGGE
FIQEDDARKLGPGRALPFPPTSECVSDVEPDTRMVRQAKKKKSGGQFQSMGL
SYVFKGIMKGVKVPPTQRTIPVLDGQDVAMARTGSGTKTACFLLPMPERLKAR
SAQTGARALILSPTELALQTKFKELCKFTGLKTALILGGDKMDQFAALHNPDI
IIATPGRLVHVAVENMLKQSVYVVDADRLFPENGPAEQEIIGRLPGGHQTVLF
SATLPKLVFARAGLTPVLRLDQVDSKINQLKTSFLLVREDTRKAVALLYLQNVV
RFQDQTVFVATKHAETYLTELLMGQGVSCAHYISALDQTKARKINAKTHNKCSTLI
VTDLAAGLDIPLLONVINYSPAKGLFLHRVGRVARGRSGTAYSLVAPDQLQGG
DLHLFGRSVTLARPCPEPSVADAVGRDGLRVQSVVDDDSSTQTAMGASLDLQGG
LHRVANNAGQVYRSPAPSPSIKAEKELDLAELGLHPLFSSCFEGELQRLVDS
IKNYRTRTTFEINASSKPPSQAKRRRMRRTORRDEFFVYPAPKPDF
LPQEEENVEVTEGVTBVGGKRPSPGQAKRRRMRRTORRDEFFVYPAPKPDF
SERGLSVGAGGAFEQQVAGAVLDLMDQANMSRQQLKMDKRRKRFVGGQGBDK
KKITSGRFFISSYKRDLYQKWKQKIDDRSEEGFSPNQRKPPRGGRGKRGQK
TSPRASSVPAGMRSELTKQILKORRQGEAALPAARPEAASTOPTPPGPAAP
GRLWPCSLQEGQDEKQVRSQTQPGWFLVAPGVDRVDYVHVLEDPSTGAAL"
2963. .2968
/polyA_signal
/notes="putative"
2984
/polyA_site
/notes="putative"
ORIGIN
Alignment Scores:
Pred. No.: 1.35 Length: 2984
Score: 129.00 Matches: 134
Percent Similarity: 31.24% Conservative: 55
Best Local Similarity: 22.15% Mismatches: 201
Query Match: 3.57% Indels: 29
DB: 3 Gaps: 29
US-09-930-591-2 (1-686) x AK031679 (1-2984)
Qy 100 CysGlySerSerAspLeuTyrLeu-----ValThrArgHisAlaAspVal 114
Db 98 TCGGAAACCCCAACTGGGGCTTCCCAAGGCGCGACACGCGACTCGATGAGCGGCAGT 157
Qy 115 IleProValArgArg-----GlyAspGlyArgGlySerLeuSerProArg 131
Db 158 TCGAGATCCAGCGCGAGGATGACGCCGCGCGAGGAAGCTGGCGCTCGCAGACCTTGC 217
Qy 132 Pro-IleSerTyrLeuLys-----GlySerSerGlyGlyProLeuLe 145
Db 218 CCTATTTCCTACCTCAGATCGGTATCAGATGTGGAGCCGACACTCGGAGATGTGTC 277
Qy 145 uCysProAlaGlyHisAlaValGlyIlePheArgAlaAlaVal----- 159
Db 278 GAGCCCAAGAACAAAGAAAGAAAGTCTGGAGGCTTCAGTCCATGGGCTGAGTTACC 337
Qy 160 -CysThr-ArgGlyValAlaLysAlaValAspPheIleProValGluSerLeuGluThrT 179
Db 338 CTGTGTTCAAGGGGATCATGAAAAGGGCTACAAAGGTGCGCGCCCATCCAGAGGAAGA 397
Qy 179 hrMetArgSerProValPheSerAspAsnSerSerProProAlaValProGlnSerTyrG 199
Db 398 CCATC-----CCCGTATCTTGGATGGCAAGGATGTGGTGGCCATG----- 438
Qy 199 lnValAlaHisLeuHisAlaProThrGlySerGlyLysSerThr----- 213
Db 439 -----GCCCGGACGAGCGAGTGGCAAGCGGCTGCTTCTCTCTCCCGA 481
Qy 214 -----LysValProAlaAlaTyrAlaAlaGlnGlyTyrLysValLeuValLeuAsnP 231
Db 482 TGTTTGAGCGGCTGAAGGACGCGAGTGCACAGACGGGGCTCGAGCCCTCATCTCTCAC 541
Qy 231 roSer-----ValAlaAlaThrMetGlyPheGlyAlaTyrMetSerLysAlaHis- 247
Db 542 CCACCGGGAGCTGGCCCTGCAGACCATGAAGTTCAATAAGAGAGTTCAGGCAAGTT-CACC 600
Qy 248 -----GlyIleAspProAsnIleArgThrGlyValArgThrIleThrGlyS 264
Db 601 GGCTCAAGACAGCCTTGTGCTGGTGGAGACAAATGGAAGACCAAGTTCAGGCTG 660

```

QY 264 erProIleThrThrTyrSerThrTyrGlyLysPheLeuAlaAspGlyGlyCysSerGlyGlyA 284  
DB : : : : :  
DB 661 CAGAGAACCTGACATATATCCACCCCTGGCGCTGTGGTGCATGT-----GGCT 714  
QY 284 laTyAspIleIleIleCysAspGluCysHisSerThrAspAlaThrSerIleLeuGlyI 304  
DB : : : : :  
DB 715 GTGGAGATGAACCTGAAGCTGCAGAGTGT-----GGAGTATGTGTG 756  
QY 304 le----- 304  
DB : :  
DB 757 TTCGATGAAGCAGACAGGCTCTTTGAATGGGCTTTGTGAGCAGCTACAGGAGATCATA 816  
QY 305 -----GlyThrValLeuAsp---GlnAlaGluThrAlaG 315  
DB 817 GGCCGCCCTTCCTGGGGCCACACAGACGGTGTCTTCTCAGCTACACTGCCAAGCTGCTG 876  
QY 315 lYAlaArgLeuThr-----ValLeuAlaThrAlaThrProProGlySerValThrV 332  
DB : : : : :  
DB 877 GTGGAAATTGCAGGGCAGGCTCACAGACCCGTCTCATCCGCTCGAGCTGAGCTAGACTCC 936  
QY 332 alProHisProAsnIleGluGluValAlaLeuSerThrThrGlyGluIlePro----- 349  
DB : : : : :  
DB 937 AAGCTCAATGAGCAGCTCAAGACCTC-----CTTCTCTCTTGTG 975  
QY 350 -----PheTyrclyLysAlaIleProLeuGluAlaIleLys-----GlyG 363  
DB 976 CGCGAAGACACCAAGGCTGTGTCTCTCTACCTGCGAGATGTGCTGGGCCCCAG 1035  
QY 363 ly-ArgHisLeuIlePheCysHisSerLysLysCysAspGluLeuAlaAlaLysLeu 382  
DB : : : : :  
DB 1036 GACACAGCTGGTGTGTGACCAAGACCAATGCGGAGTACCTCACAGAGTTGCTG 1095  
QY 383 VallAlaLeuGlyValAlaAlaValAlaTyrrTyrrArgGlyLeuAspValSerValIlePro 402  
DB : : : : :  
DB 1096 ATGGGCCAGGCTGTGAGTGGCCACATCATATAGTCTTGACACGACG----- 1146  
QY 403 ThrSerGlyAspValValValAlaThrAspAlaLeuMetThrGlyPheThrGlyAsp 422  
DB : : : : :  
DB 1147 -----GCCCGCAAGATCAACTGGCCCAAGTTCACACACAAC 1182  
QY 423 PheAspSerValIleAspCysAsnThr---CysValThrGlnThrValAspPheSerLeu 441  
DB : : : : :  
DB 1193 -----AAATGTTCCACCTCATCGTACCTGCTGCGCCGCGCGGCGCTG 1227  
QY 442 AspProThrPheThrIleGluThrIleThrLeuProGlnAspAla----- 456  
DB : : : : :  
DB 1228 GACATCCACCTGCTGACACGCTCATCACTACAGCTTCCCTGCCAAGGCAAGCTCTTC 1287  
QY 457 -----ValSerArgThrGlnArgArgGlyArgThrGlyArgGlyLysProGlyIle 473  
DB : : : : :  
DB 1288 CTGCACCGAGTGGCGCTGTGGCCGAGCAGCGCCGAGTGGCAGAGCC----- 1335  
QY 474 TyrArgPheValAlaProGlyGlu-ArgProSerGlyMetPheAspSerValLeuCy 493  
DB : : : : :  
DB 1336 TATCTTGTGGTGGCCCGACAGAGTCCCTTACC----- 1369  
QY 493 sGluCysTyrrAspAlaGlyCysAlaTyrrGluLeuThrProAlaGluThrThrValAr 513  
DB : : : : :  
DB 1370 -----TGCTTGACCTACACTGTCTCTGGCGCTCTGTACCTGCGCCGCTTGTGAG 1425  
QY 513 g---LeuArgAlaTyrrMetAsnThrProGlyLeuProValCysGlnAspHisLeuGluPh 532  
DB : : : : :  
DB 1426 GAGCTTCAGTGGCAGATGCGTGTGGCAGGG-----ACGAGGTG 1464  
QY 532 eTTPGluGly----- 535  
DB : : : : :  
DB 1465 CTGGGTCCGTCGCCCGCAGAGTGTAGTGGATGATGAGGACAGCAGCCTGCAGACTGCCATG 1524  
QY 536 -----ValPheThrGlyLeuThrHisIleAspAlaHisPheLeuSerGlnTh 551  
DB : : : : :  
DB 1525 GGGGCATCCCTGGATCTTCAGGGCTGACCC----- 1555  
QY 551 rLysGlnSerGlyGluAsnLeuProTyrrLeuValAlaTyrrGlnAlaThrValCysAlaAr 571

DB 1556 -----GGGTGGCCCAACACGCTCAGCAGCAGTATGTGGCTCA 1593  
QY 571 gAlaGlnAlaPro-----ProProSerTTPAsp---GlnMetTr 583  
DB : : : : :  
DB 1594 CGGCAGCGCCCTCGCCTGAGTCCATCAGAGAGCCAGGAGCTGGAGCTGCAGAGCTG 1653  
QY 593 pLysCys 585  
DB 1654 GGCTTGC 1660  
RESULT 14  
AK028274  
LOCUS  
DEFINITION  
Mus musculus 14, 17 days embryo head cDNA, RIKEN full-length  
enriched library, clone:322401E18 product:hypothetical P-loop  
containing nucleotide triphosphate hydrolases structure containing  
protein, full insert sequence.  
ACCESSION  
VERSION AK028274.1 GI:26080791  
KEYWORDS HTC; CAP trapper.  
SOURCE Mus musculus (house mouse)  
ORGANISM Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
1  
REFERENCE  
AUTHORS Carninci, P. and Hayashizaki, Y.  
TITLE High-efficiency full-length cDNA cloning  
JOURNAL Meth. Enzymol. 303, 19-44 (1999)  
MEDLINE 99279253  
PUBMED 10349636  
2  
REFERENCE  
AUTHORS Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,  
Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.  
TITLE Normalization and subtraction of cap-trapper-selected cDNAs to  
prepare full-length cDNA libraries for rapid discovery of new genes  
JOURNAL Genome Res. 10 (10), 1617-1630 (2000)  
MEDLINE 20499374  
PUBMED 11042159  
3  
REFERENCE  
AUTHORS Shibata, K., Itoh, M., Aizawa, K., Nagao, S., Sasaki, N., Carninci, P.,  
Konno, H., Akiyama, J., Nishi, K., Kitsuai, T., Tashiro, H., Itoh, M.,  
Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,  
Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,  
Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M.,  
Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuura, S., Kawai, J.,  
Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.  
TITLE RIKEN integrated sequence analysis (RISA) system--384-format  
sequencing pipeline with 384 multicapillary sequencer  
JOURNAL Genome Res. 10 (11), 1757-1771 (2000)  
MEDLINE 20530913  
PUBMED 11076861  
4  
REFERENCE  
AUTHORS The RIKEN Genome Exploration Research Group Phase II Team and the  
FANTOM Consortium.  
TITLE Functional annotation of a full-length mouse cDNA collection  
JOURNAL Nature 409, 685-690 (2001)  
PUBMED 11076861  
5  
REFERENCE  
AUTHORS The FANTOM Consortium and the RIKEN Genome Exploration Research  
Group Phase I & II Team.  
TITLE Analysis of the mouse transcriptome based on functional annotation  
of 60,770 full-length cDNAs  
JOURNAL Nature 420, 563-573 (2002)  
PUBMED 12042159  
6  
REFERENCE  
AUTHORS Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P.,  
Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W.,  
Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T.,  
Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kawai, I., Kasukawa, T.,  
Kato, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M.,  
Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M.,  
Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N.,  
Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N.,  
Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T.,

Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A., Muramatsu, M. and Hayashizaki, Y. Direct Submission Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration and Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Teurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gscl.riken.jp, URL: http://genome.gsc.riken.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216) cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. Please visit our web site for further details. URL: http://genome.gsc.riken.jp/ URL: http://fantom.gsc.riken.jp/ URL: http://fantom.gsc.riken.jp/.		Db	
FEATURES		394	
source		CCATC-----CCCGTGATCTTGGATGGCAGAGATGGTGGGCCATG-----	
misc_feature		21..2751 /note="hypothetical P-loop containing nucleotide triphosphate hydrolases structure containing protein (SCOP 52540, evidence: SCOP) putative"	
polyA_signal		2960..2965 /note="putative"	
polyA_site		2986 /note="putative"	
ORIGIN		332	
Alignment Scores:		350	
Pred. No.:		1.35	
Score:		129.00	
Percent Similarity:		31.24%	
Best Local Similarity:		22.15%	
Query Match:		3.57%	
DB:		3	
US-09-930-591-2 (1-686) x AK028274 (1-2986)		383	
QY		100	
Db		94	
QY		115	
Db		154	
QY		132	
Db		214	
QY		145	
Db		274	
QY		160	
Db		334	
QY		179	
Db		199	
QY		434	
QY		213	
Db		477	
QY		231	
Db		537	
QY		248	
Db		596	
QY		264	
Db		656	
QY		284	
Db		710	
QY		304	
Db		812	
QY		315	
Db		872	
QY		332	
Db		932	
QY		349	
Db		971	
QY		363	
Db		1031	
QY		382	
Db		1091	
QY		402	
Db		1142	
QY		422	
Db		1178	
QY		441	
Db		1223	
QY		456	
Db		1283	
QY		473	
Db		1331	
QY		493	
Db		1365	



QY 493 sGluCysTyrAspAlaGlyCysAlaTyrTyrGluLeuThrProAlaGluThrValAr 513  
 Db 1366 ----TGCTGACCTACACCTGTCTCGCCGCTGTACACCTCGCCGCTTGTGTGAG 1421  
 QY 513 g---LeuArgAlaTyrMetAsnThrProGlyLeuProValCysGlnAspHisLeuGluPh 532  
 Db 1422 GAGCCTTCAGTGGCAGATCGGTTGGCAGG-----ACGAGTGT 1460  
 QY 532 eTrpGluGly----- 535  
 Db 1461 CTGGGTGCGTGTGCCCCCAGAGTGTAGTGGATGATGAGGACAGCAGCCTGCAGACTGCCATG 1520  
 QY 536 -----ValPheThrGlyLeuThrHisIleAspAlaHisPheLeuSerGlnTh 551  
 Db 1521 GGGGATCCCTGGATCTTCAGGGCTGCACC----- 1551  
 QY 551 rLysGlnSerGlyGluAsnLeuProTyrLeuValAlaTyrGlnAlaThrValCysAlaAr 571  
 Db 1552 -----GCGTGCCCAACAGCTCGACGACGATGTGTGCGCTCA 1589  
 QY 571 gAlaGlnAlaPro-----ProProSerTyrAsp---GlnMetTr 583  
 Db 1590 CGGCAGCGCCTCGCCTGAGTCCATCAAGAGGCCAAGAGCTGCACCTGGCAGAGCTG 1649  
 QY 583 pLysCys 585  
 Db 1650 GCCTTGC 1656

RESULT 15  
 LOCUS BC043699 3956 bp mRNA linear HTC 13-JAN-2003  
 DEFINITION Mus musculus, Similar to RIKEN cDNA 2410015A15 gene, clone  
 IMAGE:5344158, mRNA.  
 ACCESSION BC043699  
 VERSION BC043699.1 GI:27696772  
 KEYWORDS HTC.  
 SOURCE Mus musculus (house mouse)  
 ORGANISM Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 1 (bases 1 to 3956)  
 Strausberg,R.  
 Direct Submission  
 Submitted (10-JAN-2003) National Institutes of Health, Mammalian  
 Gene Collection (MGC), Cancer Genomics Office, National Cancer  
 Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,  
 USA

REMARK  
 COMMENT NIH-MGC Project URL: <http://mgc.nci.nih.gov>  
 Contact: MGC help desk  
 Email: [cgapbs@mail.nih.gov](mailto:cgapbs@mail.nih.gov)  
 Tissue Procurement: Jeffrey Green M.D.  
 cDNA Library Preparation: Life Technologies, Inc.  
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: Genome Sequence Centre,  
 BC Cancer Agency, Vancouver, BC, Canada  
 info@bcgsc.bc.ca  
 Steven Jones, Jennifer Asano, Ian Bosdet, Yaron Butterfield,  
 Susanna Chan, Readman Chiu, Chris Fjell, Erin Garland, Ran Guin,  
 Letitia Hsiao, Martin Krzywinski, Reta Kutsche, Oliver Lee, Soo  
 Sen Lee, Victor Ling, Carrie Mathewson, Candice McLeavy, Steven  
 Ness, Pawan Pandoh, Anna-Liisa Prabhu, Parvaneh Saedi, Jacqueline  
 Schein, Duane Smalish, Michael Smith, Lorraine Spence, Jeff Stott,  
 Michael Thorne, Miranada Tsai, Natasja van den Bosch, Jill Vardy,  
 George Yang, Scott Zuyderduyn, Marco Marra.

Clone distribution: MGC clone distribution information can be found  
 through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>  
 Series: IRAC Plate: 86 Row: 1 Column: 19  
 This clone was selected for full length sequencing because it  
 passed the following selection criteria: Hexamer frequency ORF  
 analysis

This clone has the following problem: frame shifted.

FEATURES  
source

Location/Qualifiers  
 1..3956  
 /organism="Mus musculus"  
 /mol\_type="mRNA"  
 /strain="FVB/N"  
 /db\_xref="taxon:10090"  
 /clone="IMAGE:5344158"  
 /tissue\_type="Mammary tumor. C3(1)-Tag model. Infiltrating  
 ductal carcinoma. 5 month old virgin mouse."  
 /clone\_lib="NCI CGAP\_Mam6"  
 /lab\_host="DH10B"  
 /note="vector: pCMV-SPORT6"

## ORIGIN

Alignment Scores:  
 Pred. No.: 2.09 Length: 3956  
 Score: 129.00 Matches: 134  
 Percent Similarity: 31.24% Conservative: 55  
 Best Local Similarity: 22.15% Mismatches: 216  
 Query Match: 3.57% Indels: 201  
 DB: 3 Gaps: 29  
 US-09-930-591-2 (1-686) x BC043699 (1-3956)  
 QY 100 CysGlySerSerAspLeuTyrLeu-----ValThrArgHisAlaAspVal 114  
 Db 80 TCGGAAACGCCGAACCTGGGCTTCCCAAGCCGCGACAGCGACTCGGATGACGGGAGT 139  
 QY 115 IleProValArgArgArg-----GlyAspGlyArgGlySerLeuLeuSerProArg 131  
 Db 140 TCGAGATCCAGCGCGGAGGATGACGCCGGCGGAGAGCTGGGCCCTGGCAGACCTTGC 199  
 QY 132 Pro-IleSerTyrLeuLys-----GlySerSerGlyGlyProLeuLe 145  
 Db 200 CCTATTCTCTACCTCAGATGCGTATCAGATGTGGAGCCGACACTCGGAGATGGTGC 259  
 QY 145 uCysProAlaGlyHisAlaValGlyIlePheArgAlaAlaVal----- 159  
 Db 260 GAGCCCAACAAGAAAAAGAGTCTGAGGCTTCCAGTCCATGCCCTGAGTTACC 319  
 QY 160 -CysThr-ArgGlyValAlaLysAlaValAppPheIleProValGluSerLeuGluThrT 179  
 Db 320 CTGTGTTCAAGGGGATCATGAAAAAGGGCTACAAGGTGCCAGCCCATCCAGAGGAAGA 379  
 QY 179 hrMetArgSerProValPheSerAspAsnSerSerProProAlaValProGlnSerTyrG 199  
 Db 380 CCATC-----CCGTGATCTTGGATGGCAAGGATGGTGGCCATG----- 420  
 QY 199 lnValAlaHisLeuHisAlaProThrGlySerGlyLysSerThr----- 213  
 Db 421 -----GCCCGGACAGGCGAGTGGCAAGAGCGGCTGCTTCTCTCCCGA 463  
 QY 214 -----LysValProAlaAlaTyrAlaAlaGlnGlyTyrLysValLeuValLeuAnp 231  
 Db 464 TGTGTGACGCGCTGACAGCGCAGTGCACAGACGGGGCTCGAGCCCTCATCTCTCAC 523  
 QY 231 roSer-----ValAlaAlaThrMetGlyPheGlyAlaTyrMetSerLysAlaHisG 248  
 Db 524 CCACCGGGAGCTGGCCCTGCAGACCATGAGTTCACTAAAGAGCTAGGCAAGTT-CACC 582  
 QY 248 ly-----IleAspProAsnIleArgThrGlyValArgThrIleThrThrGlys 264  
 Db 583 GGCCTCAAGACTGCTTGTATCTGGTGGAGACAAATGGAAGACCACTTTGACGCCCTG 642  
 QY 264 erProIleThrTyrSerThrTyrGlyLysPheLeuAlaAspGlyCysSerGlyGlyA 284  
 Db 643 CACGAGAACCTGACATAATCATTCATCCACCCCTCGGCGCTCTGGTGCATGT-----GGCT 696  
 QY 284 laTyrAspIleIleCysAspGluCysHisSerThrAspAlaThrSerIleLeuGlyI 304  
 Db 697 GTGAGATGAACCTGAAGCTGCAGAGTGT-----GGAGTATGTGTGTG 738  
 QY 304 le----- 304

Mon Feb 28 09:44:43 2005

us-09-930-591-2.1st

Db 1636 GGCTTGC 1642

Search completed: February 27, 2005, 07:18:59  
Job time : 4768 secs

```

739  TTCCATGAGCAGACAGAGCTCTTTGAAATGGGCTTGTGAGCAGCTACAGAGATCATA 798
305  -----GlyThrValLeuAsp---GlnAlaGluThrAlaG 315
799  GGCGGCTTCTGGGGCCACAGAGGCTGCTGTTCTCAGCTACACTGCCCAAGCTGCTG 858
315  lYAlaArgLeuThr-----ValLeuAlaThrAlaThrProProGlySerValThrV 332
859  GTGGAATTCACAGCGGCGAGCTCAGAGAGCCGCTCATCGCTGAGCTGAGACTCC 918
332  alProHisProAsnIleGluGluValAlaLeuSerThrThrGlyGluIlePro----- 349
919  AAGCTCAATGAGCAGCTCAAGACCT-----CTTCCTCTCTGTG 957
350  -----PheTyrGlyLeuAlaIleProLeuGluAlaIleLys-----GlyG 363
958  CGCAAGACACACAGGCTGCTGCTCTCTACTGCTGAGAATGCTGCTGGGCCCAAG 1017
363  ly-ArgHisLeuIlePheCysHisSerLysLysCysAspGluLeuAlaLysLeu 382
1018  GACCAGACTGTGTGTTGTAGCCACAAGCACCATCGGAGTACCTCACAGAGTGTGCTG 1077
383  ValAlaLeuGlyValAsnAlaValAlaTyrTyrArgGlyLeuAspValSerValIlePro 402
1078  ATGGCCAGGCTGTGAGTGGGCCACATCTATAGTGCCTTGGACACAGC----- 1128
403  ThrSerGlyAspValValAlaThrAspAlaLeuMetThrGlyPheThrGlyAsp 422
1129  -----GCCGCAAGATCAACTTGGCCAAAGTTTCACACACAAC 1164
423  PheAspSerValIleAspCysAsnThr---CysValThrGlnThrValAspPheSerLeu 441
1165  -----AAATGTTCCACCTCATCTGACTGACCTGGCGCGCCGGGGCTG 1209
442  AspProThrPheThrIleGluThrIleThrLeuProGlnAspAla----- 456
1210  GACATCCCACTGCTGACACAGTCACTCACTCCCTGCGCAAGGCAAGCTCTTC 1269
457  -----ValSerArgThrGlnArgArgGlyArgThrGlyArgGlyLysProGlyLe 473
1270  CTGCACCGAGTGGCGCGTGTGGCCGAGCAGCGCGAAGTGGCAGCC----- 1317
474  TyrArgPheValAlaProGlyGlu-ArgProSerGlyMetPheAspSerSerValLeuCy 493
1318  TATTCTTGTGGCCCGCAGACAGAGTCCCTACC----- 1351
493  sGluCysTyrAspAlaGlyCysAlaTyrTyrGluLeuThrProAlaGluThrThrValAr 513
1352  ----TGTGACCTACACTGTTCTCTGGCGCGCTCTGTACCCCTGGCCCGCTCTGTGAG 1407
513  g---LeuArgAlaTyrMetAsnThrProGlyLeuProValCysGlnAspHisLeuGluPh 532
1408  GAGCCTTCAGTGGCAGATCGGTTGGCAGG-----ACGAGTG 1446
532  eTrpGluGly----- 535
1447  CTGGGTGCGCTGCGCCAGAGTGTAGTGGATGATGAGCAGCAGCAGCTGCCATG 1506
536  -----ValPheThrGlyLeuThrHisIleAspAlaHisPheLeuSerGlnTh 551
1507  GGGGCATCCCTGGATCTTCAGGCGCTGCACC----- 1537
551  rLysGlnSerGlyGluAsnLeuProTyrLeuValAlaTyrGlnAlaThrValCysAlaAr 571
1538  -----CGTGGCCACAACGCTCAGCAGCAGTATGTGGCTCA 1575
571  gAlaGlnAlaPro-----ProProSerTrpAsp---GlnMetTr 583
1576  CGGCCAGCGCTCTCGCTGAGTCCATCAAGAGAGCCAGGAGCTGGACCTGGCAGAGCTG 1635
583  pLysCys 585

```

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 27, 2005, 15:23:02 ; Search time 8681 Seconds  
(without alignments)  
11503.997 Million cell updates/sec

Title: US-09-930-591-1  
Perfect score: 2061  
Sequence: 1 atggcgccctacgcgcta.....atgaaatgaagagtgtga 2061

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0  
Searched: 4708233 seqs, 24227607955 residues  
Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : GenEmbl.\*  
1: gb.ba.\*  
2: gb.htg.\*  
3: gb.in.\*  
4: gb.om.\*  
5: gb.ov.\*  
6: gb.pat.\*  
7: gb.ph.\*  
8: gb.pl.\*  
9: gb.pr.\*  
10: gb.ro.\*  
11: gb.sts.\*  
12: gb.sy.\*  
13: gb.un.\*  
14: gb.vi.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	2061	100.0	2061	6	CQ826998	Sequence
2	2061	100.0	2061	6	AX441176	Sequence
3	2061	100.0	2061	6	AX467113	Sequence
4	1854.8	90.0	9610	14	HEC278830	Sequence
5	1846.4	89.6	8791	14	AY615798	Sequence
6	1787	86.7	6299	6	AX164584	Sequence
7	1786	86.7	5360	6	I06434	Sequence
8	1786	86.7	5360	6	I09328	Sequence
9	1786	86.7	6785	6	I06440	Sequence
10	1786	86.7	7310	6	AR118696	Sequence
11	1786	86.7	7310	6	I09331	Sequence
12	1786	86.7	7310	14	HPCPOLYP	Sequence
13	1786	86.7	9185	6	I08294	Sequence
14	1786	86.7	9185	6	BD091382	Sequence
15	1786	86.7	9379	6	AR166930	Sequence
16	1786	86.7	9379	6	AR301300	Sequence
17	1786	86.7	9401	6	AR176483	Sequence
18	1786	86.7	9401	6	E66593	Sequence
19	1786	86.7	9401	6	I71894	Sequence

20	1786	86.7	9401	6	I81885	I81885 Sequence 9
21	1786	86.7	9401	6	BD080334	BD080334 Hepatitis C
22	1786	86.7	9401	14	HPCPOLYP	M62321 Hepatitis C
23	1784.4	86.6	6785	6	AR118692	AR118692 Sequence
24	1784.4	86.6	6785	6	I09329	I09329 Sequence 10
25	1784.4	86.6	8316	6	AR118703	AR118703 Sequence
26	1784.4	86.6	8987	6	AR118728	AR118728 Sequence
27	1784.4	86.6	9185	6	AR118722	AR118722 Sequence
28	1784.4	86.6	9185	6	AR118723	AR118723 Sequence
29	1784.4	86.6	9379	6	AR118747	AR118747 Sequence
30	1781.2	86.4	2058	6	AR404933	AR404933 Sequence
31	1781.2	86.4	2058	6	AX395309	AX395309 Sequence
32	1781.2	86.4	5360	6	AR118686	AR118686 Sequence
33	1781.2	86.4	9424	14	AF511948	AF511948 Hepatitis
34	1781.2	86.4	9609	12	AF387805	AF387805 Synthetic
35	1781.2	86.4	9618	14	AF271632	AF271632 Hepatitis
36	1781.2	86.4	9646	12	AF387806	AF387806 Synthetic
37	1781.2	86.4	9693	12	AF387807	AF387807 Synthetic
38	1779.6	86.3	2058	6	AR408362	AR408362 Sequence
39	1779.6	86.3	2058	6	AX454818	AX454818 Sequence
40	1774.8	86.1	12980	6	AR110831	AR110831 Sequence
41	1774.8	86.1	12980	6	BD069985	BD069985 Functiona
42	1773.2	86.0	9609	12	AF387808	AF387808 Synthetic
43	1773.2	86.0	9646	6	AR110828	AR110828 Sequence
44	1773.2	86.0	9646	6	BD069982	BD069982 Functiona
45	1773.2	86.0	9646	14	AF009606	AF009606 Hepatitis

ALIGNMENTS

RESULT 1						
CQ826998						
LOCUS	CQ826998	2061 bp	DNA	linear	PAT 29-JUN-2004	
DEFINITION	Sequence 1 from Patent WO2004048402.					
ACCESSION	CQ826998					
VERSION	CQ826998.1	GI:49455655				
KEYWORDS	synthetic construct					
SOURCE	synthetic construct					
ORGANISM	synthetic construct					
REFERENCE	1					
AUTHORS	Sallberg,M.					
TITLE	A hepatitis C virus codon optimized non-structural ns3/4a fusion					
JOURNAL	Gene					
FEATURES	Patent: WO 2004048402-A 1 10-JUN-2004; TRIPEP AB (SE)					
source	Location/Qualifiers					
	1..2061					
	/organism="synthetic construct"					
	/mol_type="unassigned DNA"					
	/db_xref="taxon:32630"					
	/note="Hepatitis C virus NS3/4A coding region"					

ORIGIN

Query Match	100.0%	Score 2061;	DB 6;	Length 2061;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 2061;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1	ATGGCGCCTTATCAGCGCCTATGCCAGCAGACAGAGGGGCTTTTGGGATGCATAATCACC	60	
Db	1	ATGGCGCCTTATCAGCGCCTATGCCAGCAGACAGAGGGGCTTTTGGGATGCATAATCACC	60	
QY	61	AGCTTGACCGCGCGGACAAAACACAGGTGGAGGTTCAGATCGTGCACTGCT	120	
Db	61	AGCTTGACCGCGCGGACAAAACACAGGTGGAGGTTCAGATCGTGCACTGCT	120	
QY	121	GCCGACACTTCTGGCACTTCGATTAACGGGGTGTGTGACTCTCTACCATGAGCC	180	
Db	121	GCCGACACTTCTGGCACTTCGATTAACGGGGTGTGTGACTCTCTACCATGAGCC	180	
QY	181	GGAACAAGGACCATTCGTCACCTAAGGGTCTCTGTTATCCAGATGTACCAATGTGGAC	240	

181 GGAAACAAGGACCAATTGCGTCACTAAGGGTCTCTTTATCCAGATGTACCAATGTGGAC 240  
241 CAAGACCTGTAGGCTGGCCCGCTCCCAAGGTGCGGCTCATTAACACATGCACTTGC 300  
241 CAAGACCTGTAGGCTGGCCCGCTCCCAAGGTGCGGCTCATTAACACATGCACTTGC 300  
301 GGCTCTCGGACCTTTACCTGTGTCAGAGCAGCGCGATGTCAATCTCTGTGGCCGACGG 360  
301 GGCTCTCGGACCTTTACCTGTGTCAGAGCAGCGCGATGTCAATCTCTGTGGCCGACGG 360  
361 GGTGATGGCAGGCGAGCGTCTTTGCGCCCGGCTTATCTCTTACTTTGAAGGCTCTCG 420  
421 GGAGGCCCTCTGCTGTGCCCCCGCAGACATGCGGTAGGCATATTACAGCCCGGGTATGC 480  
421 GGAGGCCCTCTGCTGTGCCCCCGCAGACATGCGGTAGGCATATTACAGCCCGGGTATGC 480  
481 ACCCGTGGAGTGGCTAAGCGGTGGACTTTCATCCCGTAGAGCTTTAGAGCAACCATG 540  
481 ACCCGTGGAGTGGCTAAGCGGTGGACTTTCATCCCGTAGAGCTTTAGAGCAACCATG 540  
541 AGGTCCCCGGTGTCTCAGACAACTCTCCCAACAGCAGTGCCTCCAGAGCTACCAAGT 600  
541 AGGTCCCCGGTGTCTCAGACAACTCTCCCAACAGCAGTGCCTCCAGAGCTACCAAGT 600  
601 GCCACCTGCTGCTCCACCGCAGCGGTAGAGCAACCAAGTCCCGCCGCGATACGCA 660  
601 GCCACCTGCTGCTCCACCGCAGCGGTAGAGCAACCAAGTCCCGCCGCGATACGCA 660  
661 GCTCAGGGCTACAAGGTGCTGGTGTCAACCCCTCGTTGCTGCAACAATGGCTTTGTT 720  
661 GCTCAGGGCTACAAGGTGCTGGTGTCAACCCCTCGTTGCTGCAACAATGGCTTTGTT 720  
721 GCTTACATGTCCAAGGCCCATGGGATTTGATCTCTAAATCAGGAGTGGGGTGAGCAATT 780  
721 GCTTACATGTCCAAGGCCCATGGGATTTGATCTCTAAATCAGGAGTGGGGTGAGCAATT 780  
781 ACTACTGGCAGCCGATACGATATTCACCTAGGCAAGTCTCTGCGCACCGCGGTGT 840  
781 ACTACTGGCAGCCGATACGATATTCACCTAGGCAAGTCTCTGCGCACCGCGGTGT 840  
841 TCAGGGGGTGTATTGACATAATAATTTGTGACGAGTGCCACTCCACCGATGCAACATCC 900  
841 TCAGGGGGTGTATTGACATAATAATTTGTGACGAGTGCCACTCCACCGATGCAACATCC 900  
901 ATCTTGGGATTTGGCACTGTCTTGTGACCAAGCAGAGACCGCGGGGCGAGCTGACTGTG 960  
901 ATCTTGGGATTTGGCACTGTCTTGTGACCAAGCAGAGACCGCGGGGCGAGCTGACTGTG 960  
961 CTGCGCACCGCTACCCCTCGGGCTCGGTCACTGTGCCCCCATCTTAACATCGAGAGGTT 1020  
961 CTGCGCACCGCTACCCCTCGGGCTCGGTCACTGTGCCCCCATCTTAACATCGAGAGGTT 1020  
1021 GCTCTGTCCACTACCGAGAGATCCCTTTTATGGCAAGCTATTCCCTCTGAAGCAATT 1080  
1021 GCTCTGTCCACTACCGAGAGATCCCTTTTATGGCAAGCTATTCCCTCTGAAGCAATT 1080  
1081 AAGGGGGGAGACATCTCATCTTCTGCACTCAAGAAAGAGTGCGACGAGCTCGCCGCA 1140  
1081 AAGGGGGGAGACATCTCATCTTCTGCACTCAAGAAAGAGTGCGACGAGCTCGCCGCA 1140  
1141 AAATGGTTCGGTGGGCTCAATGCGGTGGCTTACTACCGGGCTTATGATGTCCGTC 1200  
1141 AAATGGTTCGGTGGGCTCAATGCGGTGGCTTACTACCGGGCTTATGATGTCCGTC 1200  
1201 ATCCCCACAGTGGTGACCTGTGTCGTCGGAACCTGACGCCCTCATATGACCGCTTTAC 1260  
1201 ATCCCCACAGTGGTGACCTGTGTCGTCGTCGGAACCTGACGCCCTCATATGACCGCTTTAC 1260  
1261 GCGCACTTCGATTCGGTGTAGACTGCAACACGCTGTGTCAACCGAGACAGCTGCACTTACG 1320  
1261 GCGCACTTCGATTCGGTGTAGACTGCAACACGCTGTGTCAACCGAGACAGCTGCACTTACG 1320

1321 CTTGACCTTACCTTCAACATTTGAGACAATCAGCTTCCCGAGATGCTGCTCTCCGTA 1380  
1321 CTTGACCTTACCTTCAACATTTGAGACAATCAGCTTCCCGAGATGCTGCTCTCCGTA 1380  
1381 CAACCTCGGGGTAGGACTGGCAGAGGGAAGCCAGGCATCTACAGATTTTGTGGCACCGGG 1440  
1381 CAACCTCGGGGTAGGACTGGCAGAGGGAAGCCAGGCATCTACAGATTTTGTGGCACCGGG 1440  
1441 GAGCGTCTTCTGGCATGTTTGAATCTGCTGCTCTCTGCGAGTGTATGACGCGGTTGT 1500  
1441 GAGCGTCTTCTGGCATGTTTGAATCTGCTGCTCTCTGCGAGTGTATGACGCGGTTGT 1500  
1501 GCTTGGTATGAGCTTACGCGCCCGCAGACCAAGTTAGGCTACGAGCATACATGAACACC 1560  
1501 GCTTGGTATGAGCTTACGCGCCCGCAGACCAAGTTAGGCTACGAGCATACATGAACACC 1560  
1561 CCGGACTTCCCGTGTGCAAGCAACCTTTGAATTTTGGAGGGGCTCTTTACGGGTCTC 1620  
1561 CCGGACTTCCCGTGTGCAAGCAACCTTTGAATTTTGGAGGGGCTCTTTACGGGTCTC 1620  
1621 ACCCATAGAGCGCCACTTCTATCCAGACAAGAGAGTGGGGAACCACTTCCCTAT 1680  
1621 ACCCATAGAGCGCCACTTCTATCCAGACAAGAGAGTGGGGAACCACTTCCCTAT 1680  
1681 CTGTAGCTGACCAAGCCACCGTGTGCGTACAGCTCAAGCCCTCCCGCTCGTGGGAC 1740  
1681 CTGTAGCTGACCAAGCCACCGTGTGCGTACAGCTCAAGCCCTCCCGCTCGTGGGAC 1740  
1741 CAGATGTGGAAGTGTGATCCGCTCTCAAGCCCACTCCATGGGCAACACCTCTGCTA 1800  
1741 CAGATGTGGAAGTGTGATCCGCTCTCAAGCCCACTCCATGGGCAACACCTCTGCTA 1800  
1801 TATAGACTGGGCGCTGTCCAGAAATGAAGTCACTGACGACCCAGTCAACCAAGTATATC 1860  
1801 TATAGACTGGGCGCTGTCCAGAAATGAAGTCACTGACGACCCAGTCAACCAAGTATATC 1860  
1861 ATGACATGTATGTGCGGTGACCTGAGGTGCTGACAGTACCTGCTGCTGCTGGCGG 1920  
1861 ATGACATGTATGTGCGGTGACCTGAGGTGCTGACAGTACCTGCTGCTGCTGGCGG 1920  
1921 GTTCTGGCTGCTTTGGCGCGTATTGCTTATCCATGACAGGTGCTGCTGCTAGTAGG 1980  
1921 GTTCTGGCTGCTTTGGCGCGTATTGCTTATCCACAGGCTGCTGCTGCTAGTAGG 1980  
1981 ATTGCTTGTCCGGAAGCGGCAATCATACCCGACAGGAAGTCTCTACCGGAGTTC 2040  
1981 ATTGCTTGTCCGGAAGCGGCAATCATACCCGACAGGAAGTCTCTCTACCGGAGTTC 2040  
2041 GATGAATGGAAGAGTGTGA 2061  
2041 GATGAATGGAAGAGTGTGA 2061

## RESULT 2

AX441176  
LOCUS AX441176 2061 bp DNA linear PAT 28-JUN-2002  
DEFINITION Sequence 16 from Patent WO0213855.  
ACCESSION AX441176  
VERSION AX441176.1 GI:21665758  
KEYWORDS synthetic construct  
SOURCE synthetic construct  
ORGANISM other sequences; artificial sequences.  
REFERENCE 1  
AUTHORS Sallberg, M. and Hultgren, C.  
TITLE Vaccines containing ribavirin and methods of use thereof  
JOURNAL Patent: WO 0213855-A 16 21-FEB-2002;  
TRIPEP AB (SE)  
FEATURES  
Location/Qualifiers  
1..2061  
/organism="synthetic construct"  
/mol\_type="unassigned DNA"

/db_xref="taxon:32630" /note="Hepatitis C virus NS3/4A coding region"									
ORIGIN	Query Match	100.0%;	Score 2061;	DB 6;	Length 2061;				
	Best Local Similarity	100.0%;	Pred. No. 0;	Mismatches	0;	Indels	0;	Gaps	0;
	Matches 2061;	Conservative	0;						
QY	1	ATGGGCGCTATCACGGCCCTATGCCAGCAGACAAAGGGGCTTTTGGGATGCATAATCAC	60						
DB	1	ATGGGCGCTATCACGGCCCTATGCCAGCAGACAAAGGGGCTTTTGGGATGCATAATCAC	60						
QY	61	AGCTTGACCGCGGAGCAAAACCAAGGTGAGAGGTGAGGTTCAAGTCTGTCAACTGCT	120						
DB	61	AGCTTGACCGCGGAGCAAAACCAAGGTGAGAGGTGAGGTTCAAGTCTGTCAACTGCT	120						
QY	121	GCCAGACTTCTTGGCAACCTGCATTAAACGGGGTGTGAGCTGTGACATGCTTACCATGGAGCC	180						
DB	121	GCCAGACTTCTTGGCAACCTGCATTAAACGGGGTGTGAGCTGTGACATGCTTACCATGGAGCC	180						
QY	181	GGAAACAAGGACCAATTCGCTCACTAAGGGTCTCTGTTATCCAGATGTACCAATGTGGAC	240						
DB	181	GGAAACAAGGACCAATTCGCTCACTAAGGGTCTCTGTTATCCAGATGTACCAATGTGGAC	240						
QY	241	CAAGACCTCGTAGGGTGGCCCGCTCCCAAGGTGCCCGTCTCATTTAAACCATGCACTTGC	300						
DB	241	CAAGACCTCGTAGGGTGGCCCGCTCCCAAGGTGCCCGTCTCATTTAAACCATGCACTTGC	300						
QY	301	GGCTCTCGGACTTTACCTGTGTCAGAGGACGCGGATGTCATTCCTGTGCGCGACGG	360						
DB	301	GGCTCTCGGACTTTACCTGTGTCAGAGGACGCGGATGTCATTCCTGTGCGCGACGG	360						
QY	361	GGTGATGGCAGGGGACGCTTCTCGCCCGGCTTATCTTTACTTTGAAAGGCTCTCG	420						
DB	361	GGTGATGGCAGGGGACGCTTCTCGCCCGGCTTATCTTTACTTTGAAAGGCTCTCG	420						
QY	421	GGAGGCGCTCTGTGTGTCGCGCAGACATGCGGTAGGCATATTCAGAGCGCGGTATGC	480						
DB	421	GGAGGCGCTCTGTGTGTCGCGCAGACATGCGGTAGGCATATTCAGAGCGCGGTATGC	480						
QY	481	ACCGTGGAGTGGCTAAGCGGTGAGCTTCATCCCGGTAGAGGCTTAGAGCAACCATG	540						
DB	481	ACCGTGGAGTGGCTAAGCGGTGAGCTTCATCCCGGTAGAGGCTTAGAGCAACCATG	540						
QY	541	AGGTCCCGGCTTCTCAGACAACTCTCCCAACAGCAGTGCCTCCAGAGCTACCAAGTG	600						
DB	541	AGGTCCCGGCTTCTCAGACAACTCTCCCAACAGCAGTGCCTCCAGAGCTACCAAGTG	600						
QY	601	GCCACCTGCTGCTCCCAACCGGCTAAGAGCAACCAAGTCCCGCGCGATACGCA	660						
DB	601	GCCACCTGCTGCTCCCAACCGGCTAAGAGCAACCAAGTCCCGCGCGATACGCA	660						
QY	661	GCTCAGGGCTACAGGTGCTGGTCTCAACCCCTCGTTGCTGCAACATGGGCTTGGT	720						
DB	661	GCTCAGGGCTACAGGTGCTGGTCTCAACCCCTCGTTGCTGCAACATGGGCTTGGT	720						
QY	721	GCTTACATGTCCAAAGGCCCATGGGATTCCTTAACATCAGGACTGGGTGAGGACAAAT	780						
DB	721	GCTTACATGTCCAAAGGCCCATGGGATTCCTTAACATCAGGACTGGGTGAGGACAAAT	780						
QY	781	ACTACTGGCAGCCGATCAAGTATTCCTACCGCAAGTTCCTCCCAACCGCGGTGT	840						
DB	781	ACTACTGGCAGCCGATCAAGTATTCCTACCGCAAGTTCCTCCCAACCGCGGTGT	840						
QY	841	TCAGGGGGTCTTATGACATAAATTTGTGAGAGTGCCTCCAGAGTGCACATCC	900						
DB	841	TCAGGGGGTCTTATGACATAAATTTGTGAGAGTGCCTCCAGAGTGCACATCC	900						
QY	901	ATCTTGGGCAATTCGCACTGTCTTGACCAAGCAGAGACCGCGGGGCGGACTGACTGTG	960						
DB	901	ATCTTGGGCAATTCGCACTGTCTTGACCAAGCAGAGACCGCGGGGCGGACTGACTGTG	960						
QY	961	CTGCCACCGCTACCCCTCGGGCTCCGTCACGTGTGCCCCATCTTAACATCGAGGAGTT	1020						
DB	961	CTGCCACCGCTACCCCTCGGGCTCCGTCACGTGTGCCCCATCTTAACATCGAGGAGTT	1020						
QY	1021	GCTCTGTCCACTACCGGAGAGATCCCTTTTATGTGCAAGGCTATTCCCTTGAAGCAAT	1080						
DB	1021	GCTCTGTCCACTACCGGAGAGATCCCTTTTATGTGCAAGGCTATTCCCTTGAAGCAAT	1080						
QY	1081	AAGGGGGGAGACATCTCATCTTCTGCGCATCAAAAGAAAGTGGAGCGAGCTCGCGCA	1140						
DB	1081	AAGGGGGGAGACATCTCATCTTCTGCGCATCAAAAGAAAGTGGAGCGAGCTCGCGCA	1140						
QY	1141	AAACTGCTGCGGTGGGCGTCAATCGCGGTCTACTACCGGGGCTTGAATGTCCGTC	1200						
DB	1141	AAACTGCTGCGGTGGGCGTCAATCGCGGTCTACTACCGGGGCTTGAATGTCCGTC	1200						
QY	1201	ATCCCGACCAAGTGTGAGCTTGTCTGTGGGCAACTGACCGCCCTCATGACCGGCTTACC	1260						
DB	1201	ATCCCGACCAAGTGTGAGCTTGTCTGTGGGCAACTGACCGCCCTCATGACCGGCTTACC	1260						
QY	1261	GGCGACTTTCGATTCGATAGCTGCAACACGTTGTGTCACCGACAGTTCGACTTTCAGC	1320						
DB	1261	GGCGACTTTCGATTCGATAGCTGCAACACGTTGTGTCACCGACAGTTCGACTTTCAGC	1320						
QY	1321	CTTGACCTTACCTTCAACATTCGAGCAATCACGTTCCCGAGATGCTCTCCCGTACT	1380						
DB	1321	CTTGACCTTACCTTCAACATTCGAGCAATCACGTTCCCGAGATGCTCTCCCGTACT	1380						
QY	1381	CAACGCTCGGGTAGGACTGGCAGAGGAGCCAGCATCTACAGATTTCCTGGCACCGGG	1440						
DB	1381	CAACGCTCGGGTAGGACTGGCAGAGGAGCCAGCATCTACAGATTTCCTGGCACCGGG	1440						
QY	1441	GAGCGCTTCTTGGCATGTTGACTGCTGTCCTCTGCGAGTGTATGACGCGGTTGT	1500						
DB	1441	GAGCGCTTCTTGGCATGTTGACTGCTGTCCTCTGCGAGTGTATGACGCGGTTGT	1500						
QY	1501	GCTTGGTATGAGCTTACGCGCGCGCAGACCAAGTTAGCTACGAGCATACATGACACC	1560						
DB	1501	GCTTGGTATGAGCTTACGCGCGCGCAGACCAAGTTAGCTACGAGCATACATGACACC	1560						
QY	1561	CGCGGACTTCCCGTGTGCGCAAGCATCTTGAATTTTGGAGGGGCTCTTTACCGGCTCTC	1620						
DB	1561	CGCGGACTTCCCGTGTGCGCAAGCATCTTGAATTTTGGAGGGGCTCTTTACCGGCTCTC	1620						
QY	1621	ACCCACATAGAGCCCACTTCTTCCAGACAAAGCAGAGTGGGAAAACTTCCCTAT	1680						
DB	1621	ACCCACATAGAGCCCACTTCTTCCAGACAAAGCAGAGTGGGAAAACTTCCCTAT	1680						
QY	1681	CTGGTAGCTACCAAGCACCGTGTGCGCTAGAGCTCAAGCCCTCCCGGCTGGGAC	1740						
DB	1681	CTGGTAGCTACCAAGCACCGTGTGCGCTAGAGCTCAAGCCCTCCCGGCTGGGAC	1740						
QY	1741	CAGATGTGGAAGTGTGATCCGTCTCAAGCCCACTTCCATGGGGCAACACCTCTGTCTA	1800						
DB	1741	CAGATGTGGAAGTGTGATCCGTCTCAAGCCCACTTCCATGGGGCAACACCTCTGTCTA	1800						
QY	1801	TATAGACTGGGCGCTGTCCAGAAATGAAGTCAACCTGACGACCCAGTCAACAGTATATC	1860						
DB	1801	TATAGACTGGGCGCTGTCCAGAAATGAAGTCAACCTGACGACCCAGTCAACAGTATATC	1860						
QY	1861	ATGACATGTATGTGGCTGACCTGGAGGTGCTCAAGTACCTGGGTGCTCTGTGGGCGC	1920						
DB	1861	ATGACATGTATGTGGCTGACCTGGAGGTGCTCAAGTACCTGGGTGCTCTGTGGGCGC	1920						
QY	1921	GTTCTGGCTGCTTGGCGGCTATTGCTTATCCAGAGGTGCTGCTCATAGTAGTAGG	1980						
DB	1921	GTTCTGGCTGCTTGGCGGCTATTGCTTATCCAGAGGTGCTGCTCATAGTAGTAGG	1980						
QY	1981	ATTGTCTTGTCCGAAAGCCGCAATCATACCCGACAGGAAAGTCTCTACCGGGAGTTTC	2040						
DB	1981	ATTGTCTTGTCCGAAAGCCGCAATCATACCCGACAGGAAAGTCTCTACCGGGAGTTTC	2040						
QY	2041	GATGAAATGGAAAGTGTCTGA	2061						







/product="non-structural protein 5b"  
/note="ORF10"

ORIGIN			
Query Match 90.0%; Score 1854.8; DB 14; Length 9610;			
Best Local Similarity 93.8%; Pred. No. 0;			
Matches 1931; Conservative 0; Mismatches 127; Indels 0; Gaps 0;			
QY	2	TGCGCGCTATCACGGCTATGCCAGACAGCAAGGGGCTTTTGGGATGCATAATCACCA	61
DB	3418	TGCGGCGCATACGGGGTACGCCAGACAGCAAGGGGCTTTTGGGATGCATAATCACCA	3477
QY	62	GCTTGACCGGCGGGCAAAAACACAGGTGGAGGGTGAGGTTGAGATCGTCAACTGCTG	121
DB	3478	GCCTGACCGGCGGGCAAAAACACAGGTGGAGGGTGAGGTTGAGATCGTCAACTGCTG	3537
QY	122	CCGAGACTTCTTGGCAACTGCAATTAACGGGGTGTTGGAGTGTCTACCATGGAGCGG	181
DB	3538	CCGAGACTTCTTGGCAACTGCAATTAACGGGGTGTTGGAGTGTCTACCATGGAGCGG	3597
QY	182	GAACAAGGACCAATTGCGTCACCTTAAGGGTCTCTTATCCAGATGTACCAATGTGGACC	241
DB	3598	GAACAAGGACCAATGCGTCACCTTAAGGGTCTCTTATCCAGATGTACCAATGTAGACC	3657
QY	242	AAGACCTCGTAGGCTGGCCCGCTCCCAAGGTGCCCGCTCATTAACACCATGCACTTGG	301
DB	3658	AAGACCTCGTAGGCTGGCCCGCTCCCAAGGTGCCCGCTCATTAACACCTGCACTTGG	3717
QY	302	GCTCCTCGAGCTTACCTGTCACGAGGACGCGCATGTCACTCTGTGGCGGACGGG	361
DB	3718	GCTCCTCGAGCTTACCTGTCACGAGGACGCGCATGTCACTCTGTGGCGGACGGG	3777
QY	362	GTGATGGAGGGGACGCTGCTTTCGCCCGGCTATCTCTTACCTTGAAGGCTCCTCGG	421
DB	3778	GTGATGGAGGGGACGCTGCTTTCGCCCGGCTATCTCTTACCTTGAAGGCTCCTCGG	3837
QY	422	GAGGCGCTCTGTGTGCCCGGAGGACATGCCGTAGGCATATTCAGAGCGCGGTATGCA	481
DB	3838	GAGGCGCTCTGTGTGCCCGGAGGACATGCCGTAGGCATATTCAGAGCGCGGTATGTA	3897
QY	482	CCCGTGGAGTGTCTAAGGCGGTGGACTTCATCCCGTAGAGCTTTAGAGACAACCATGA	541
DB	3898	CCCGTGGAGTGTCTAAGGCGGTGGAAATTTGTCTCTGTAGAGAACCTTAGAGACAACCATGA	3957
QY	542	GGTCCCGGCTGTCTCAGACAACTCTCCCGACAGCAGTCCCGAGAGCTTACCAAGTGG	601
DB	3958	GGTCCCGGCTGTCTCAGACAACTCTCCCGACAGCAGTCCCGAGAGCTTCCAGGTGG	4017
QY	602	CCCACTGTATGCTCCACCGGCGGTAAAGACACCAAGTCCCGGCGCATACGCAG	661
DB	4018	CCCACTGTATGCTCCACCGGCGGTAAAGACACCAAGTCCCGGCGCTTACGGCG	4077
QY	662	CTCAGGGCTACAAGGTGCTGGTCTCAACCCCTCGTTGCTGCTCAACAATGGGCTTTGGT	721
DB	4078	CTCAGGGCTACAAGGTGCTGGTCTCAACCCCTCGTTGCTGCTCAACAATGGGCTTTGGT	4137
QY	722	CTTACATGTCGAAGGCCATGGATTGATCTTAACATCAGAGCTGGGGTGAGACAATTA	781
DB	4138	CTTATATGTCGAAGGCCATGGATTGATCTTAACATCAGAGCTGGGGTGAGACAATTA	4197
QY	782	CTACTGGACCGCGATCAGTATTCACCTACGGCAAGTTCCTTCCGACCGCGGGTGT	841
DB	4198	CCACCGGACGCGCATCAGTACTCCACTTACGGCAAGTTCCTTCCGACCGCGGGTGT	4257
QY	842	CAGGGGTCTTTATGACATAAATTTGTGAGAGTGGCCACTCCACGGATGCAATCCA	901
DB	4258	CAGGGGTCTTTATGACATAAATTTGTGAGAGTGGCCACTCCACGGATGCAATCCA	4317
QY	902	TCTTGGGCACTGGCACTGCTCTTGAACAGAGAGACCGCGGGGCGAGACTGACTGTGC	961
DB	4318	TCTTGGGCACTGGCACTGCTCTTGAACAGAGAGACCGCGGGGCGAGGCTGGTTGTC	4377
QY	962	TCGCCACCGCTACCCCTCCGGGCTCCGTCACTGTGCCCATCTTAACATCGAGGAGGTG	1021

DB	4378	TCGCCACCGCTACCCCTCCGGGCTCCGTCACTGTGCCCATCTTAACATCGAAGAGTTG	4437
QY	1022	CTCTGTCCACTACCGGAGGATCCCTTTTATGGCAAGGCTATTCCTTGAAGCAATTA	1081
DB	4438	CTCTGTCCACTACCGGAGGATTCCTTTTATGGCAAGGCTATTCCTTGAAGCAATTA	4497
QY	1082	AGGGGGGAGAGCATCTCATCTTCTGCCACTCAAGAAGAGTGCGACAGCTCGCGCAA	1141
DB	4498	AGGGGGGAGAGCATCTCATCTTCTGCCACTCAAGAAGAGTGCGACAGCTCGCGCAA	4557
QY	1142	AACTGGTCGGTGGGCGTCAATGGCGTGTACTACCGCGGCTTGTGTGTCGTCA	1201
DB	4558	AACTGGTCGGTGGGCGTCAATGGCGTGTACTACCGCGGCTTGTGTGTCGTCA	4617
QY	1202	TCCGACCAAGTGGTCAGCTTGTGCTGGGCACTGACGCCCTCATGACCGGCTTACCG	1261
DB	4618	TCCGACCAAGTGGTCAGCTTGTGCTGGGCACTGACGCCCTCATGACCGGCTTACCG	4677
QY	1262	GCGACTTCGATTCGGTGTAGTACTGCAACACGCTGTGTCAACCAGAGTGCATTCAGCC	1321
DB	4678	GCGACTTCGATTCGGTGTAGTACTGCAACACGCTGTGTCAACCAGAGTGCATTCAGCC	4737
QY	1322	TTGACCTTCACTTCACTTACCATTTGAGACAATCAAGCTTCCCAGGATGCTGTCTCCGCTAC	1381
DB	4738	TTGACCTTCACTTCACTTACCATTTGAGACAATCAAGCTTCCCAGGATGCTGTCTCCGCTAC	4797
QY	1382	RACGTGGGGTAGGACTGGCAGAGGAAGCAAGGATCTACAGATTTGTGGCACCGGGG	1441
DB	4798	RACGTGGGGTAGGACTGGCAGAGGAAGCAAGGATCTACAGATTTGTGGCACCGGGG	4857
QY	1442	AGCGTCTTCTGGCATGTTTGACTGCTGCTCTGCTGAGTGTATGACGCGGCTGTG	1501
DB	4858	AGCGTCTTCTGGCATGTTTGACTGCTGCTCTGCTGAGTGTATGACGCGGCTGTG	4917
QY	1502	CTTGTATGAGCTTACGCCCGGAGACCAAGTTAGCTACGAGCATACATGAACACCC	1561
DB	4918	CTTGTATGAGCTTACGCCCGGAGACCAAGTTAGCTACGAGCATACATGAACACCC	4977
QY	1562	CGGGACTTCCCGGTGCGCAAGCCATCTTGAATTTTGGAGGGGCTCTTACGGGCTCTCA	1621
DB	4978	CGGGACTTCCCGGTGCGCAAGCCATCTTGAATTTTGGAGGGGCTCTTACGGGCTCTCA	5037
QY	1622	CCCAATAGAGCGCCACTTCTTATCCAGACAAAGCAGAGTGGGGAACCTTCCCTATC	1681
DB	5038	CCCAATAGAGCGCTCACTTCTTATCCAGACAAAGCAGAGTGGGGAACCTTCTTACC	5097
QY	1682	TGGTAGCGTACCAAGCCACCGTGTGCGTAGAGCTCAAGGCCCTCCCGCTGCTGGGACC	1741
DB	5098	TGGTAGCGTACCAAGCCACCGTGTGCGTAGAGCTCAAGGCCCTCCCGCTGCTGGGACC	5157
QY	1742	AGATGTGGAAGTGTGATCCGTCTCAAGCCACCTCCATGGGCGCAACCTCTGCTAT	1801
DB	5158	AGATGTGGAAGTGTGATCCGTCTCAAGCCACCTCCATGGGCGCAACCTCTGCTAT	5217
QY	1802	ATAGACTGGGCGCTGCTCCAGAAATGAAGTCACTCCAGCACCCAGCTACCAAGTATATCA	1861
DB	5218	ATAGACTGGGCGCTGCTCCAGAAATGAAGTCACTCCAGCACCCAGCTACCAAGTATATCA	5277
QY	1862	TGACATGATGTGCGGCTGACCTGGAGGTGCTCAGAGTACCTTGGGTGCTGCTGGCGG	1921
DB	5278	TGACATGATGTGCGGCTGATCTGGAGATCGTCAAGTACCTGGGTGCTGCTGGCGG	5337
QY	1922	TTCTGGCTGCTTGGCGCGGATTTGCCCTATCCAGGCTGGTGTGCTATAGTATAGTA	1981
DB	5338	TTCTGGCTGCTTGGCGCGGATTTGCCCTATCCAGGCTGGTGTGCTATAGTATAGTA	5397
QY	1982	TTGCTTCTCGGGAAGCGGCAATCATACCCGACAGGGAAGTCTCTTACCGGAGTTCG	2041
DB	5398	TTGCTTCTCGGGAAGCGGCAATCATACCCGACAGGGAAGTCTCTTACCGGAGTTCG	5457
QY	2042	ATGAATGGAAGAGTGCT 2059	

Db 5458 ATGAGATGGAAGAGTGCT 5475

RESULT 5  
AY615798  
LOCUS  
DEFINITION  
Hepatitis C virus isolate HCV.TWB.1 polyprotein gene, partial cds.  
ACCESSION  
AY615798  
VERSION  
AY615798.1 GI:48479029  
KEYWORDS  
Hepatitis C virus  
SOURCE  
Hepatitis C virus  
ORGANISM  
Hepatitis C virus  
VIRUSES; serona positive-strand viruses, no DNA stage; Flaviviridae; Hepacivirus.

REFERENCE  
1 (bases 1 to 8791)  
AUTHORS  
Brann, T.W., Kottlilil, S., Polis, M. and Imamichi, T.  
TITLE  
Identification of mutations associated with interferon resistance in HCV and HIV co-infected patients  
JOURNAL  
Unpublished  
2 (bases 1 to 8791)  
AUTHORS  
Brann, T.W., Kottlilil, S., Polis, M. and Imamichi, T.  
TITLE  
Direct Submission  
JOURNAL  
Submitted (03-MAY-2004) LHR/CSP, SAIC-Frederick, Inc, Building 550, Room 126, 1050 Boyles Street, Frederick, MD 21702, USA

FEATURES  
source  
1..8791  
/organism="Hepatitis C virus"  
/mol\_type="genomic RNA"  
/strain="HCV1a"  
/isolate="HCV.TWB.1"  
/db\_xref="taxon:11103"  
/country="USA"  
67..>8791  
/codon\_start=1  
/product="polyprotein"  
/protein\_id="A414436.1"  
/db\_xref="GI:48479030"

CDS  
translation="MSTNPKPQKRTKRNRRPDQVKKFGGGQIVGGVYLLPRRGRRL  
GVRAKTSERSQPRGRORIPKAREPEGRWAQPGYWPVLYNEGCGWAGWLLSPRG  
SRPQWPTPRRRSLKVIDTLCGFADLMGYIPLVAGPLGAARALAHGVRLVED  
GVNATNLPQGSFSI FLAALLSCLTPASAYOVNRNSGLYHVNDPCNSIIVETAD  
AILHSGCPCVREGNASKCWA VAVTATRDGRLPSTQLRHIDLLVGSATLCSALY  
VGLCGSVLQGLQTFSPRRHWTQDCNSIYPGHITGRHAWDMNMWSTTALV  
AQLLRTPAILDIAGAHGVLGAIAYFSMVGNKAVLVLLFAGVDAETSTTGAA  
ARTAYFTSFAPGPQNVQIYIYNGSWHINRTALNCASLDITGMIALFYNSFNS  
CGPERMASCRLPASPDQGPISYANGSGPEHRPYCWHYPKPGI VPAQSGVPYC  
FTSPVVGTTDKFQVPTYNWANDTDVFLANNTRPLGNWFGCTWNSSTGTKVCGA  
PPCIVGVGNTHLCTDCPRKHPEATYSRCGSPWITPRCLVDYFYRLWHYPTLNY  
TLFKRMVYGVGVEHREAAACNWTGRERCDDDRSELSPLLLTTTQWQVLPSCFTIL  
PALTTGLIHLQNIIVDQYLVGVGSSI VSWAIKWEYVILLFLLADARICSLMMWLL  
ISOAEALLENLI VLNAAISLAGTHGLASFVFFCFAWYLGKWPVGNMAYAFPGWPLLL  
LLIALPQRAVALDTEMAASCGLVVLGLMALTLSPHYKRYISMCLWMLQVFLTRABAL  
LQWVPLNVRGGRDAVILLMCVHPALVPDITKLIALALGPLWILQTSLLKVPYFVR  
VQGLIRICALARKAGHIVQMAIIKAGALTGYIYNHITPLDMHNGRLDLAVAVE  
PVFVSQMETKLTITWADTAACGDI INGLPVARSRRREIILGPADGNASKWGRLLAPIT  
AYAQOETRGILGCIITLGRDNQVGEI QIVSTAQTFLATCINGCVWTVHGAOTR  
TIASPKGPVOMYNTIDQLVGWPAQGRASLPCTCGSSDLVLTVRHADVIPVRERG  
DSRGLSPRII SYLKGSSGGLLCPAGHAGVIFRAACTRGVAKAVDFIPVENLETT  
MRSPVTDNSPRAVQSGFOVAHLHAPTSGSKTKVPAAYAGQYKVLNPSVAATL  
GFGAYMSKAGHDNPTITGVRTITIGSPITITSGRLADGGCGGAYDIIICDECHS  
TDATSLIGTVDLQDAETAGARLVLATATPGSITVPHPNIEVALSTSLTFPFYK  
AI PLEAI KGRHDIIFCHSKKCDLAKLVALGINAVAYYRGDLDSVPTSGDVVVA  
TDALMTQGTDFIDNCVTQTDVDFSLDPTFTIETTLPODAVSRTQRRGTRGRG  
KPGIYRFA PGERPSGMFSSVLCYDAGCANVELT PAETTVRLRAYMNTPLGVQO  
DHLFWEQVETGLTIDAHFSLQTKSGENLPVLYAYQATVCARAQAPPPSWDMWKC  
LIRLKPTLHGPTLLRLVLAGVQNEVLTIPVTKIIMTCSADLEVSTSVTLVGGVLA  
ALAAAYCLSTGCWVIGRI VLSGRPAVIPDRELYRFDENEECQHLPIYEQMMLAE  
QYKQALGLIQASROAEVIA PAVQTNWQKLEAFWAKHMMNFISGQIYLAGLSTLEN  
PAIASLMAFTAVTSLTTSQTLLENILGSWAAQLAAGAAAFVAGLAGLAGIAGV  
GLKVLVDIILAGYAGVAGLVAFKIMSGEPPTEDLVNLPAILSPGALVGVVCA  
ILRHVGPBGAGVQWNRLLIAPASRGNHVSPTHVYVESDAARVTAIISLTVXLLR  
RLHWVSSECTTPCSGSWLRDIIDWICELVSLDITWILKAKMLPQLGPIPVSCQRGR  
GVNRGDGIMHTRCHCGABITGHVRKNGTMRITVGPKTCRNWMSGTFPFINAYTGPCTPLP

ORIGIN

Query Match 89.6%; Score 1846.4; DB 14; Length 8791;  
Best Local Similarity 93.5%; Pred. No. 0;  
Matches 1925; Conservative 1; Mismatches 132; Indels 0; Gaps 0;

QY 2 TGGCGCCTATCAGCGCTATGCCAGACAGCAAGGGCGCTTTTGGGATCATATAATCACCA 61  
DB 3143 TGGCGCCTATCAGCGCTATGCCAGACAGCAAGGGCGCTTTAGGATGATATAATCACCA 3202

QY 62 GCTTGACCGCGCGGACAAAAACACAGGTGGAGGTGAGTTTCTGATCGTGTCACTGCTG 121  
DB 3203 GTCTGACCGCGCGGACAAAAACACAGGTGGAGGTGAGATTGATGATGATGCTG 3262

QY 122 CCCAGACTTTCTTGGCAACTGCTATTAACGGGCTGTGTGGACTGTCTACCATGGAGCGG 181  
DB 3263 CCCAGACTTTCTTGGCAACTGCTATTAACGGGCTGTGTGGACTGTCTACCATGGAGCGG 3322

QY 182 GAACAAGGACCATTTGCTGCTCACTAAGGGTCTGTATCCAGATGTACCAATGTGGACC 241  
DB 3323 GAACAAGGACCATTTGCTGCTCACTAAGGGTCTGTATCCAGATGTATATAATAGACC 3382

QY 242 AAGACTCTAGGCTGGCGCGCTCCCAAGAGTGCCTCATTAACCATGACCTTGG 301  
DB 3383 AAGACTCTAGGCTGGCGCGCTCCCAAGAGTGCCTCATTTGACGCCCTGACCTTGG 3442

QY 302 GCTCTCTCGGACCTTTTACCTGTGTCAGGACGACCGCATGTCTCTCTGTCGCGCGGCGG 361  
DB 3443 GCTCTCTCGGACCTTTTACCTGTGTCAGGACGACCGCATGTCTCTCTGTCGCGCGGCGG 3502

QY 362 GTGATGGCAGGGGACGCTGCTTTTCCGCCCGGCTATCTCTTACTTTAAAGGCTCTCTCG 421  
DB 3503 GTGATGGCAGGGGACGCTGCTTTTCCGCCCGGCTATCTCTTACTTTAAAGGCTCTCTCG 3562

QY 422 GAGGCGCTCTGCTGTCGCCCGGACGATGCGGTAGGATATTCAGAGCGCGCGGTATGCA 481  
DB 3563 GAGGCGCTCTGCTGTCGCCCGGACGATGCGGTAGGATATTCAGAGCGCGCGGTATGCA 3622

QY 482 CCGGTGAGTGGCTAAGGCGGTGACTTCTATCCCGTATAGAGCTTATAGACAAACATGA 541  
DB 3623 CCGGTGAGTGGCTAAGGCGGTGACTTCTATCCCGTATAGAGCTTATAGACAAACATGA 3682

QY 542 GGTCCCCGGTGTCTCAGACAACTCTCTCCCGACACAGCAGTGCCTCCAGAGCTACCAAGTGG 601  
DB 3683 GGTCCCCGGTGTCTCAGACAACTCTCTCCCGACACAGCAGTGCCTCCAGAGCTTCCAGTGG 3742

QY 602 CCCACTGCTATGCTCCCAACCGGCGGTGAAGACACCAAGGTCCTCCCGCGGTATGCGAG 661  
DB 3743 CCCACTGCTATGCTCCCAACCGGCGGTGAAGACACCAAGGTCCTCCCGCGGTATGCGAG 3802

QY 662 CTCAGGGCTACAGGCTGCTGCTCAACCCCTCCCTGCTGCAACATGGGCTTTGGTG 721  
DB 3803 CCCAGGGCTACAGGCTGCTGCTCAACCCCTCCCTGCTGCAACATGGGCTTTGGTG 3862

QY 722 CTTTACATGTCGAAGGCCCATGGATTGATCTTAACATCAGACTGGGGTGGAGCAATTA 781  
DB 3863 CTTTATATGTCGAAGGCCCATGGATTGATCTTAACATCAGACTGGGGTGGAGCAATTA 3922







Db 2971 ATGAGATGGAAGAGTGCT 2988

RESULT 8  
109328  
LOCUS 109328 5360 bp DNA linear PAT 02-DEC-1994  
DEFINITION Sequence 8 from Patent WO 8904669.  
ACCESSION 109328  
VERSION 109328.1 GI:587963  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 5360)  
AUTHORS Houghton, M., Choo, Q.-K. and Kuo, G.  
JOURNAL Patent: WO 8904669-A 8 01-JUN-1989;  
FEATURES  
source  
1..5360  
/organism="unknown"  
/mol\_type="unassigned DNA"

ORIGIN

Query Match 86.7%; Score 1786; DB 6; Length 5360;  
Best Local Similarity 91.7%; Pred. No. 0;  
Matches 1888; Conservative 0; Mismatches 170; Indels 0; Gaps 0;

QY 2 TGCGCGCTATCAGCGCTATGCCAGCAGACAGGGGCGCTTTGGGATGCATAATCACCA 61  
DB 931 TGGCGCCCATCAGCGGTACGCCAGCAGACAGGGGCGCTTCTAGGGTGCATATCACCA 990

QY 62 GCTTGACCGCGCGGCAAAAACAGGTGGAGGTTGAGTTTCAGATCGTGTCAACTGCTG 121  
DB 991 GCCTAACTGCGCGGACAAAACAAAGTGAGGGTGAGGTCCAGATTGTGTCAACTGCTG 1050

QY 122 CCAGACTTCTTGGCAACCTGCATTAAACGGGGTGTGGAGTGTCTACCATGGAGCGG 181  
DB 1051 CCCAAACCTTCTTGGCAACGTGCATCAATGGGGTGTGCTGAGCTGTCTACCAAGGGCGG 1110

QY 182 GAACAAGGACATTGGCTGACCTTAAGGGTCTCTGTTATCCAGATGTACCAATGTGGACC 241  
DB 1111 GAACGAGGACCATCGGTGTCACCAAGGGTCTCTGTTATCCAGATGTATCCATGTAGACC 1170

QY 242 AAGACCTCGTGGCTGCGCGCTCCCAAGGTGCGCGCTCATTAACACCATGCATTGCG 301  
DB 1171 AAGACCTTGTGGCTGCGCGCTCCCAAGGTAGCGCTCATTTGACACCTGCACTTGG 1230

QY 302 GCTCTCGGACCTTTACCTGGTCAAGGAGCAGCGGATGTCATCTGTGGCGCGAGGG 361  
DB 1231 GCTCTCGGACCTTTACCTGGTCAAGGAGCAGCGGATGTCATTTCCCGTGGCGCGGG 1290

QY 362 GTGATGCGAGGGGAGCGCTGCTGCGCGCGCGCGCTATCTTAAAGGCTCCTCGG 421  
DB 1291 GTGATGCGAGGGGAGCGCTGCTGCGCGCGCGCGCTATTTCTTAAAGGCTCCTCGG 1350

QY 422 GAGCGCTCTGTGTGCGCGCGAGGACATGCCGTAGGCATATTCAGAGCGCGCGGTATGCA 481  
DB 1351 GGGGTGCGCTGTGTGCGCGCGGCGACCGCTGGGCAATTTAGGGCGCGGTGTGCA 1410

QY 482 CCGGTGAGGTGGCTAAGCGCGGTGGATTCTATCCCGGTAGAGGCTTAGAGACAAACATGA 541  
DB 1411 CCGGTGAGGTGGCTAAGCGCGGTGGATTCTATCCCGGTGGAGAACCTAGAGAACCATGA 1470

QY 542 GGTCCCGGTTCTCAGACACTCTCCCGCGCGCGCGCGCGCGCGAGCTACCAAGTGG 601  
DB 1471 GGTCCCGGTTCTCAGACACTCTCTCTCCAGTAGTGCCCGCGAGCTTCAGGTGG 1530

QY 602 CCCACCTGCATGCTCCACCGCGAGCGGTAAAGACACCAAGGTCCCGCGCGCATACGCAG 661  
DB 1531 CTACCTCCATGCTCCACAGGAGCGGCAAGGACCAAGGTCCCGCGCTGCATATGCAG 1590

QY 662 CTCAGGGCTACAAGGTGCTGGTGTCAACCCCTCGGTGTGTGCAACATGGGCTTGGTG 721  
DB 1591 CTCAGGGCTACAAGGTGCTAGTACTCAACCCCTCTGTGTGTGCAACACTGGGCTTGGTG 1650

QY 722 TTATCATGTCCAAGGCCCATGGGATTGATCTCTAAATCATAGGACTGGGGTGAGCAAAATTA 781  
DB 1651 TTATCATGTCCAAGGGCTATGGGATCGATCTCTAAATCATAGGACCGGGGTGAGAAATTA 1710

QY 782 CTATGCGCAGCCCGATCAGGTATTCACCTACGGCAAGTTCTTCCCGAGCGGGGTGT 841  
DB 1711 CCACTGGCAGCCCGCATCAGTACTCACTACGGCAAGTTCTTCCCGAGCGGGGTGT 1770

QY 842 CAGGGGTGCTTATGACATAATAATTTGTGAGAGTGCCACTCCACGGATGCAACATCCA 901  
DB 1771 CGGGGGCGCTTATGACATAATAATTTGTGAGAGTGCCACTCCACGGATGCCACATCCA 1830

QY 902 TCTTGGGCATTGGCCTACTGCTTGAACAGAGAGACCGCGGGGGGAGACTGACTGTGTGC 961  
DB 1831 TCTTGGGCATTGGCCTACTGCTTGAACAGAGAGACTCGGGGGGAGACTGCTGTGTGC 1890

QY 962 TCGCCACCGCTACCCCTCGGGGCTCGGTCACTGTGCCCCCATCTTAACATCGAGGAGGTG 1021  
DB 1891 TCGCCACCGCCACCCCTCGGGGCTCGGTCACTGTGCCCCCATCTTAACATCGAGGAGGTG 1950

QY 1022 CTCTGTCCACTACCGGAGAGATCCCTTTTATGGCAAGGCTATTTCCCTTTGAAGCAATTA 1081  
DB 1951 CTCTGTCCACTACCGGAGAGATCCCTTTTATGGCAAGGCTATTTCCCTTTGAAGTATTA 2010

QY 1082 AGGGGGGAGAGACATCTCATCTTCTGCTCACTCAAAAGAAAGTGCAGAGCTCGCGCAA 1141  
DB 2011 AGGGGGGAGAGACATCTCATCTTCTGCTCACTCAAAAGAAAGTGCAGAGAACTCGCGCAA 2070

QY 1142 AACTGTCGCGTGGCGCTCAATGCGGTGCTTACTACCGCGGCTTGATGTGTCGCTCA 1201  
DB 2071 AGCTGTCGCTATGGGCATCAATGCGGTGCGCTTACTACCGCGGCTTGATGTGTCGCTCA 2130

QY 1202 TCCCGACCACTGTCAGCTGTCGTGCGCAACTGACGCCCTCATGACCGGCTTTACCG 1261  
DB 2131 TCCCGACCACTGTCAGCTGTCGTGCGCAACTGTCGTGCGCAACTGTCGTGTCGCTATACCG 2190

QY 1262 GCGACTTCGATTCGCTGATGAGCTGCAACACGCTGTGTCAACAGAGAGTGCAGCTTCAGCC 1321  
DB 2191 GCGACTTCGACTCGGTGATGAGCTGCAATACGCTGTGTCAACAGAGAGTGCAGCTTCAGCC 2250

QY 1322 TTGACCTTACCTTACCATTTGAGACATCAAGCTTCCCGAGGATGTGTCTCCGCTACTC 1381  
DB 2251 TTGACCTTACCTTACCATTTGAGACATCAAGCTTCCCGAGGATGTGTCTCCCGCACTC 2310

QY 1382 AACGTCGGGTAGGACTGCGAGAGGAGCAGGCAATACAGATTGTGGCAACCGGGG 1441  
DB 2311 AACGTCGGGTAGGACTGCGAGGAGGAGCAGGCAATACAGATTGTGGCAACCGGGG 2370

QY 1442 AGCGTCTCTTCTGCGCATGTTTGTACTGCTGCTCTCTGCGAGTGTATGACGCGGGTGTG 1501  
DB 2371 AGCGCCCTCTCGGCATGTTTGTACTGCTGCTCTCTGCGAGTGTATGACGCGGGTGTG 2430

QY 1502 CTTGGTATGAGCTTACGCGCGCGGAGACCAAGTTAGGCTACGAGATACATGAACACC 1561  
DB 2431 CTTGGTATGAGCTTACGCGCGCGGAGACTACAGTTAGGCTACGAGGCTACATGAACACC 2490

QY 1562 CGGGACTTCCCGTGTGCCAAGACCATCTTGAATTTTGGGAGGCGCTTTTACGGGTCTCA 1621  
DB 2491 CGGGGTCTCCCGTGTGCCAAGGACCATCTTGAATTTTGGGAGGCGCTTTTACAGGCTCA 2550

QY 1622 CCACATAGACGCCCACTTCTTATCCAGACAAAGCAGAGTGGGGAAACCTTTCCCTATC 1681  
DB 2551 CTATATAGTGCCTTCTATCCAGACAAAGCAGAGTGGGGAAACCTTTCCCTTACC 2610

QY 1682 TGTGAGGCTACCAAGCCACCGGTGTGCTAGAGCTCAAGCCCTTCCCGCTGTGGAGC 1741  
DB 2611 TGTGAGGCTACCAAGCCACCGGTGTGCTAGAGCTCAAGCCCTTCCCGCTGTGGAGC 2670

QY 1742 AGATGTGGAGTGTCTTGTGCTCAAGCCACCTTCAAGCGGCAACACCTCTGTCTAT 1801  
DB 2671 AGATGTGGAGTGTCTTGTGCTCAAGCCACCTTCAAGCGGCAACACCTCTGTCTAT 2730







```
Db 2704 CTTGGTATGAGCTACGCGCGCGAGACTACAGTTAGGCTTACGAGCGTACATGAACACCC 2763
Qy 1562 CGGGAATTCCTCCGTGTGCAAGACCATCTTGAAATTTTGGGAGGCGCTTTTACGGGTCTCA 1621
Db 2764 CGGGCTTCCCGTGTGCCAGGACCACTTGAAATTTTGGGAGGCGCTTTTACAGGCTCA 2823
Qy 1622 CCACATAGACGCCCACTTCTTATCCAGACAAAGAGAGAGTGGGAAAACCTTCCCTATC 1681
Db 2824 CTATATAGATGCCCACTTTCTATCCAGACAAAGAGAGTGGGAGAACCTTCCCTTACC 2883
Qy 1682 TGGTAGCTACCAAGCACCGCTGTGCGTAGAGCTCAAGCCCTCCCGCTGTGGGACC 1741
Db 2884 TGGTAGCTACCAAGCACCGCTGTGCGTAGAGCTCAAGCCCTCCCGCTGTGGGACC 2943
Qy 1742 AGATGTGGAAGTGTGATCCGCTCAAGCCCAACCTCCATAGGCGCAACACCTCTGCTAT 1801
Db 2944 AGATGTGGAAGTGTGATCCGCTCAAGCCCAACCTCCATAGGCGCAACACCTCTGCTAT 3003
Qy 1802 ATAGATGGGCGCTGTCCAGAAATGAAGTCAACCTGACGCAACCCAGTCAACAAATATCA 1861
Db 3004 ACAGACTGGGCGCTGTTCAGAAATGAAGTCAACCTGACGCAACCCAGTCAACAAATATCA 3063
Qy 1862 TGACATGTATGTGGGCTGACTGGAGTGTGCTACGAGTACCTGGGTGCTGTGGGCGG 1921
Db 3064 TGACATGTATGTGGGCGGCTGTGCTATCCACAGCTGCGTGTGCTGTGGGCGG 3123
Qy 1922 TTTCTGTGTCCGGAAGCCGCAATCATATCCGACAGGGAAGTCTCTACCGGAGTTGG 2041
Db 3184 TCGTCTTGTCCGGAAGCCGCAATCATATCCGACAGGGAAGTCTCTACCGGAGTTGG 3243
Qy 2042 ATGAATGGAAGTGTCT 2059
Db 3244 ATGAGATGGAAGTGTCT 3261

RESULT 10
AR118696
LOCUS AR118696 7310 bp DNA linear PAT 16-MAY-2001
DEFINITION Sequence 74 from patent US 6150087.
ACCESSION AR118696
VERSION AR118696.1 GI:14100606
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 7310)
AUTHORS Chien,D.Y.
TITLE NANBV diagnostics and vaccines
JOURNAL Patent: US 6150087-A 74 21-NOV-2000;
FEATURES Location/Qualifiers
source 1..7310
/organism="unknown"
/mol_type="unassigned DNA"

ORIGIN
Query Match 86.7%; Score 1786; DB 6; Length 7310;
Best Local Similarity 91.7%; Pred. No. 0;
Matches 1888; Conservative 0; Mismatches 170; Indels 0; Gaps 0;
Qy 2 TGGGGCTATCAGGGCTATGCCAGCAGACAGGGGCGCTTTTGGAGTGATTAATCACCA 61
Db 1729 TGGGGCTATCAGGGCTATGCCAGCAGACAGGGGCGCTTCTAGGTGTGATAATCACCA 1788
Qy 62 GCTTGACCGCGGGACAAAACAGGTGAGGGTTCAGATCGTCACTACCTGCTG 121
Db 1789 GCCTAACTGCCGGGACAAAACAGGTGAGGGTTCAGATCGTCACTACCTGCTG 1848
Qy 122 CCCAGACTTTCTTGGCAACCTGCATTAAACGGGGTGTGTGGACTGTCTACCATGGAGCGG 181
```

```
Db 1849 CCCAAACCTTCTCGCAACGTCATCAATGGGGTGTGTGGACTGTCTACCAACGGGCGG 1908
Qy 182 GAACAAGGACCATTCGGTCACCTAAGGGTCTGTTTATCCAGATGTACACCAATGTGGACC 241
Db 1909 GAACGAGGACCATTCGGTCACCTAAGGGTCTGTCATCCAGATGTATACCAATGTAGACC 1968
Qy 242 AAGACTTCGTAGGCTGGGCGGCTCCCAAGGTGCCGCTCATTAACACCATGACATTTGGG 301
Db 1969 AAGACTTCGTAGGCTGGGCGGCTCCCAAGGTAGCGCTCATTAACACCATGACATTTGGG 2028
Qy 302 GCTCTCTCGACCTTTTACCTGCTCAGAGGACGCGCATGTCTTCTCTGTGGCGGACGGG 361
Db 2029 GCTCTCTCGACCTTTTACCTGCTCAGAGGACGCGCATGTCTTCTCTGTGGCGGACGGG 2088
Qy 362 GTGATGGACAGGCGGACCTTCTTCCCGCGGCTATCTTCTTACTTTGAAAGGCTCTCTCGG 421
Db 2089 GTGATGACAGGCGGACCTTCTTCCCGCGGCTATCTTCTTACTTTGAAAGGCTCTCTCGG 2148
Qy 422 GAGGCGCTCTGTGTGCCCGGACGACATGCCGTAGGCATATTTAGAGCGCGGCTATGCA 481
Db 2149 GGGGTCCGCTGTGTGCCCGGCGGACGCGCTGGGCATATTTAGGGCGCGGCTGTGCA 2208
Qy 482 CCGCTGGAAGTGGCTAAGCGGCTGGACTTCATCCCGTAGAGAGCTTAGAGACAACCATGA 541
Db 2209 CCGCTGGAAGTGGCTAAGCGGCTGGACTTCATCCCGTAGAGAGCTTAGAGACAACCATGA 2268
Qy 542 GGTCCCGCGGTCTTCAGACAACCTCTCCCGACAGCAGTCCCGCAGAGCTTACCAAGTGG 601
Db 2269 GGTCCCGCGGTCTTCAGACAACCTCTCTCCAGCAGTGTGTCGCCAGAGCTTCCAGGTGG 2328
Qy 602 CCCACCTGCATCTCTCCCGCAGCGGTAAAGACCAAGAGTCCCGCGGACATACGCAG 661
Db 2329 CTCACCTCCATGCTCCCGCAGCGGCAAAAGCACCAGGTCCCGGCTGCATATGCAG 2388
Qy 662 CTCAGGGCTACAAGGTGTGTGTCTCAACCCCTCCGTTCCTGTGGAACAATGGGCTTTGGTG 721
Db 2389 CTCAGGGCTATAAGGTGTGTGTCTCAACCCCTCTGTCTGCAACACTGGGCTTTGGTG 2448
Qy 722 CTTTACATGTCCAAAGCCCATGGATTTGATCTTAAACATCAGGACTGGGGTGAGACAATTA 781
Db 2449 CTTTACATGTCCAAAGCCCATGGATTTGATCTTAAACATCAGGACTGGGGTGAGACAATTA 2508
Qy 782 CTACTGGCAGCCCGATCAGTATTTCCACCTACGCGAAGTTTCTTTCGCGAGCGGGGTGT 841
Db 2509 CCACCTGGCAGCCCGATCAGTATTTCCACCTACGCGAAGTTTCTTTCGCGAGCGGGGTGT 2568
Qy 842 CAGGGGGTCTTATGACATAATAATTTGTGACGAGTGCCTCCACGAGTGCACAACTCCA 901
Db 2569 CAGGGGGCGCTTATGACATAATAATTTGTGACGAGTGCCTCCACGAGTGCACAACTCCA 2628
Qy 902 TCTTGGGCAATTGGCACTGTCTTGACAGAGAGACCGGGGGGAGACTGACCTGTGC 961
Db 2629 TCTTGGGCAATCGGCACTGTCTTGACAGAGAGACTCGGGGGGAGACTGTTGTGTGC 2688
Qy 962 TCGCCACCGCTACCCCTCCCGGCTCCGTCACTGTGCCCATCTTAACATCGAGAGGTTG 1021
Db 2689 TCGCCACCGCTACCCCTCCCGGCTCCGTCACTGTGCCCATCTTAACATCGAGGAGTTG 2748
Qy 1022 CTCTGTCCAATAACCGGAGAGATCCCTTTTATGGCAAGGCTATTTCCCTTTAGCAATTA 1081
Db 2749 CTCTGTCCAATAACCGGAGAGATCCCTTTTATGGCAAGGCTATTTCCCTTTAGCAATTA 2808
Qy 1082 AGGGGGGAGAGATCTCATCTTCTGCGCACTCAAGAGAGTGGCGAGCTCGCGGCA 1141
Db 2809 AGGGGGGAGAGATCTCATCTTCTGCTCAATTCATTCAGAGAGTGGCGAGCTCGCGGCA 2868
Qy 1142 AACTGTCTCGGCTGGCGCTCAATCGCGGCTTACTTACCGCGCTTGTGTGTGCTGCTCA 1201
Db 2869 AACTGTCTCGGCTGGCGCTCAATCGCGGCTTACTTACCGCGCTTGTGTGTGCTGCTCA 2928
Qy 1202 TCCCGACCAAGTGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1261
```

```

Db 2929 TCCGACACAGCGCGATGTTGTCGTGTCGGCAACGATGCCCTCATGACCGGTATACCG 2988
Qy 1262 GCGACTTCGATTCGGGTAGACTGCAACACGCTGTGTACCCAGACAGTCGATTCAGCC 1321
Db 2989 GCGACTTCGATTCGGGTAGACTGCAACACGCTGTGTACCCAGACAGTCGATTCAGCC 3048
Qy 1322 TTGACCTTACCTTACCTTACCTTACCTTACCTTACCTTACCTTACCTTACCTTACCTT 1381
Db 3049 TTGACCTTACCTTACCTTACCTTACCTTACCTTACCTTACCTTACCTTACCTTACCTT 3108
Qy 1382 AACCTCGGGGTAGACTGCGAGGGAAGCCAGGCAATCTACAGATTTGTGGACCGGGGG 1441
Db 3109 AACCTCGGGGTAGACTGCGAGGGAAGCCAGGCAATCTACAGATTTGTGGACCGGGGG 3168
Qy 1442 AGCTCTCTTCGCGATTTGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1501
Db 3169 AGCGCCCTTCGCGATTTGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3228
Qy 1502 CTTGTATGAGTTACGCCCGCGAGACCACTTGAATTTTGGAGGCGTCTTTACGGGTCTCA 1561
Db 3229 CTTGTATGAGTTACGCCCGCGAGACTTACAGTTAGGCTACGAGCGTACATGAACACCC 3288
Qy 1562 CGGAGCTTCGCTGTGCGCAAGACCACTTGAATTTTGGAGGCGTCTTTACGGGTCTCA 1621
Db 3289 CGGAGCTTCGCTGTGCGCAAGACCACTTGAATTTTGGAGGCGTCTTTACAGGCGCTCA 3348
Qy 1622 CCAATAGACGCCCACTTCTCTATCCAGACAAAGCAGAGTGGGAAACCTTCCCTATC 1681
Db 3349 CTCATATAGTGCCTTCTCTATCCAGACAAAGCAGAGTGGGAAACCTTCTCTATC 3408
Qy 1682 TGGTAGCTTACCAAGCCACCGTGTGCGTGTAGAGTCAAGCCCTTCCCGCTGTGGGACC 1741
Db 3409 TGGTAGCTTACCAAGCCACCGTGTGCGTGTAGAGTCAAGCCCTTCCCGCTGTGGGACC 3468
Qy 1742 AGATGTGGAAGTGTGATTCGCTTCAAGCCACCGTGTAGAGTCAAGCCCTTCCCGCTGT 1801
Db 3469 AGATGTGGAAGTGTGATTCGCTTCAAGCCACCGTGTAGAGTCAAGCCCTTCCCGCTGT 3528
Qy 1802 ATAGACTGGCGGTGTCCAGATGAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAG 1861
Db 3529 ACAGACTGGCGGTGTTCAGATGAATCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAG 3588
Qy 1862 TGACATGTATGTGCGGTACCTTGGAGTGTGTCAAGTACCTTGGGTGTGTGTGTGTGTGTGT 1921
Db 3589 TGACATGTATGTGCGGTACCTTGGAGTGTGTCAAGTACCTTGGGTGTGTGTGTGTGTGTGT 3648
Qy 1922 TTTGCTGCTTTTGGCGCGTATTCGCTATCCAGAGGCTGCTGTGTGTGTGTGTGTGTGTGT 1981
Db 3649 TCCCTGGCTTTTGGCGCGTATTCGCTATCCAGAGGCTGCTGTGTGTGTGTGTGTGTGTGT 3708
Qy 1982 TTTGCTTGTCCGAAAGCCGCAATCATACCCGACAGGGAAGTCTCTTACCGGAGTTCCG 2041
Db 3709 TCGTCTTGTCCGAAAGCCGCAATCATACCTTACCTTACCTTACCTTACCTTACCTTACCTT 3768
Qy 2042 ATGAATGGAAGAGTGTCT 2059
Db 3769 ATGAGATGGAAGAGTGTCT 3786

```

RESULT 11

LOCUS I09331 7310 bp DNA linear PAT 02-DEC-1994  
 DEFINITION Sequence 15 from Patent WO 890469.

ACCESSION I09331

VERSION I09331.1 GI:587966

KEYWORDS

SOURCE Unknown.

ORGANISM Unclassified.

REFERENCE 1 (bases 1 to 7310)

AUTHORS Houghton, M., Choo, Q.-K. and Kuo, G.

JOURNAL Patent: WO 890469-A 15 01-JUN-1989;

FEATURES Location/Qualifiers

```

source 1..7310
/organism="unknown"
/mol_type="unassigned DNA"

ORIGIN

Query Match 86.7%; Score 1786; DB 6; Length 7310;
Best Local Similarity 91.7%; Pred. No. 0;
Matches 1888; Conservative 0; Mismatches 170; Indels 0; Gaps 0;

Qy 2 TGGCGCCCTATCAGCGCCTATGCCAGACAGCAAGGGGGCTTTTGGATGCAATAACACCA 61
Db 1729 TGGCGCCCTATCAGCGCCTATGCCAGACAGCAAGGGGGCTTCTAGGGTGCATAAATCACCA 1788
Qy 62 GCTTACACCGCGCGGACAAACCAAGGTGGAGGGTGCAGATTCAGATTCGATTCGATTCGAT 121
Db 1789 GCTTAACTGGCGCGGACAAACCAAGGTGGAGGGTGCAGATTCGATTCGATTCGATTCG 1848
Qy 122 CCCAGACTTTCTTGGCAACCTGCATTAACGGGGTGTGTGGACTGTCTACCAATGAGAGCGG 181
Db 1849 CCCAAACCTTCTGCGCAACGTGCATCAATGGGGTGTGTGGACTGTCTACCAACGGGGCG 1908
Qy 182 GAACAAGACCATTCGCTCACCTAAGGGTCCCTGTTATCCAGATGTACACCAATGTGGACC 241
Db 1909 GAACGAGGACCATTCGCTCACCAAGGGTCCCTGTTATCCAGATGTATACCAATGTAGACC 1968
Qy 242 AAGACCTCTGCTAGGCTGGCGCCGCTCCCAAGGTGCCCGCTCATTAACACCATGCTTGGC 301
Db 1969 AAGACCTTGTGGCTGGCGCCGCTCCCAAGGTAGCGCTCATTTGACACCTGCTTGGC 2028
Qy 302 GCTTCTCGGACCTTTTACCTGTGTACAGGACAGCGCGATGTCTTCTGTGCGCCGACGGG 361
Db 2029 GCTTCTCGGACCTTTTACCTGTGTACAGGACAGCGCGATGTCTTCTGTGCGCCGCGG 2088
Qy 362 GTGATGGCAGGGGACGCTGCTTTCGCGCCCGGCTTCTTCTTCTTCTTCTTCTTCTTCT 421
Db 2089 GTGATGGCAGGGGACGCTGCTTTCGCGCCCGGCTTCTTCTTCTTCTTCTTCTTCTTCT 2148
Qy 422 GAGGCGCTCTGCTGTGTGCGCGCAGGACATGCGCTAGGCAATATTCAGAGCCGCGTATGCA 481
Db 2149 GGGGTCCGCTGTGTGTGCGCGCAGGACATGCGCTAGGCAATATTTAGGCGCGCGTGTGCA 2208
Qy 482 CCGCTGGAGTGGCTTAAGCGGTGTGATTCATCCCGCTAGAGAGCTTACAGACACATGA 541
Db 2209 CCGCTGGAGTGGCTTAAGCGGTGTGATTCATCCCGCTAGAGAGCTTACAGACACATGA 2268
Qy 542 GGTCCCGCGTGTCTCAGACAACTCTCCACACAGAGTGCCTCCAGAGCTTACCAAGTGG 601
Db 2269 GGTCCCGCGTGTCTCAGAGTAACTCTCTCCACAGTGTGCTCCAGAGCTTCCAGGTGG 2328
Qy 602 CCCACCTGCATGCTCCCGCAGCGAGCTTAAGAGCAACCAAGGTCCCGCGCGCATACGCAG 661
Db 2329 CTCACCTCCATGCTCCCGCAGCGAGCTTAAGAGCAACCAAGGTCCCGCGCGCATATGCAG 2388
Qy 662 CTCAGGGCTACAAGGTGTGTGTCTCAACCCCTCCGTTGTGTGTGTGTGTGTGTGTGTGT 721
Db 2389 CTCAGGGCTATAAAGGTGTGTGTCTCAACCCCTCTCTGTGTGTGTGTGTGTGTGTGTGT 2448
Qy 722 CTTATCATGTCCAAGGCCCATGAGATTGATCCTTAACATCAGGACTGCGGGTGAACAATTA 781
Db 2449 CTTATCATGTCCAAGGCCCATGAGATTGATCCTTAACATCAGGACTGCGGGTGAACAATTA 2508
Qy 782 TTACTTGGCAGCCCGCATCAGTATTCACCTACGCGCAAGTTCCTTCCCGAGCGCGGTGTGT 841
Db 2509 CCACCTGGCAGCCCGCATCAGTATTCACCTACGCGCAAGTTCCTTCCCGAGCGCGGTGTGT 2568
Qy 842 CAGGGGGTGTGTATGACATTAATAATTTGTGAGAGTGTGTGTGTGTGTGTGTGTGTGTGT 901
Db 2569 CAGGGGGTGTGTATGACATTAATAATTTGTGAGAGTGTGTGTGTGTGTGTGTGTGTGTGT 2628
Qy 902 TCTTGGGCTATGGGCACTGTCTTGAACAGAGACAGCGCGGGGGGAGAGCTGACTGTGTGC 961
Db 2629 TCTTGGGCTATGGGCACTGTCTTGAACAGAGAGAGCTGCGGGGGGAGAGCTGACTGTGTGC 2688

```



TRCFDSTVTESDIRTEAIIYQCDLDPOARVAIKSLITERLYVGGPLTNSRGNGCVRR  
CRASGVLTSCGNLTICYIKARAACRAAGLODCTMLVCGDDLVWICESAGVODASL  
RAFTEAMTYSAPBDDPOPEVDLELLTSCSNVSVAHDGAKRXYVLTTRDPTPLAR  
AAWETARHPVNSWLNINIFAPTLWARMILMTHFVSVLIAKDQLEQALDCBIYACY  
SIEPLDLPIIQLR"

ORIGIN

Query Match	86.7%;	Score 1786;	DB 14;	Length 7310;
Best Local Similarity	91.7%;	Pred. No. 0;		
Matches 1888;	Conservative 0;	Mismatches 170;	Indels 0;	Gaps 0;
QY	2	TGCGCGCTATACGGCCTATGCCAGCAGACAAGGGCCCTTTGGGATGCATAATCACCA	61	
Db	1729	TGGCGCCCATCACGGCGTAGCGCCAGCAGACAAGGGCCCTCTAGGGTGCAATACCA	1788	
QY	62	GCTTGACCGCCGGGCAAAAACACAGGTGGAGGGTGAGTTTCAGATCGTCAACTGCTG	121	
Db	1789	GCTTAACGTGGCGGGCAAAAACCAAGTGGAGGGTGAGTTCAGATTTGTCAACTGCTG	1848	
QY	122	CCCAGACTTTCTGGCAACTGCAATTAACGGGTGTGGTGTCTACCATGGAGCCG	181	
Db	1849	CCCAACCTTCTGGCAACTGTCATCAATGGGTGTGGTGTCTACCATGGAGCCG	1908	
QY	182	GAACAAGGACCAATTCGTCACTAAGGGTCTGTATTCCAGATGTACCAATGTGGACC	241	
Db	1909	GAACGAGGACCAATTCGTCACTAAGGGTCTGTATTCCAGATGTATACCAATGTAGACC	1968	
QY	242	AAGACCTGTAGCTGGCGGCTCCCAAGGTGCCCGCTCATTACACCATGCTTGGC	301	
Db	1969	AAGACCTGTGGGCTGGCGGCTCCGCAAGGTAGCCGCTCATTTGACACCCCTGCCTGG	2028	
QY	302	GCTCTCCGACCTTTACCTGGTCAACGAGGACGCGCATGTCTTCCGTGGCCGACGGG	361	
Db	2029	GCTCTCCGACCTTTACCTGGTCAACGAGGACGCGCATGTCTTCCGTGGCCGCGGG	2088	
QY	362	GTGATGGCAGGGCAGCTGCTTTTCCCGCCGCTATCTCTTACTTTGAAAGCTCCTCGG	421	
Db	2089	GTGATAGCAGGGCAGCTGCTGCTGCGCCCGGCCCATTTCTTACTTTGAAAGCTCCTCGG	2148	
QY	422	GAGGCCCTGTGTGTCGCGGACGAGCATGCGGTAGGCATATTTCAGAGCGCGGTATGCA	481	
Db	2149	GGGGTCCGTGTGTGTGTCGCGGCGGACGCGGTGGGCATTTTAGGCGCGCGGTGTGCA	2208	
QY	482	CCCGTGGAGTGGCTAAGGCGGTGGACTTTCATCCCGGTAGAGACTTAGAGACAACCATGA	541	
Db	2209	CCCGTGGAGTGGCTAAGGCGGTGGACTTTTATCCCTGTGAGAACCTTAGAGACAACCATGA	2268	
QY	542	GGTCCCGGTGTCTCAGACAACCTCTCCCAACAGCAGTGCCTCCAGAGCTACCAAGTGG	601	
Db	2269	GGTCCCGGTGTCTCAGGATAACTCTCTCCACAGTAGTGCCTCCAGAGCTTCCAGGTGG	2328	
QY	602	CCCACCTGCATGCTCCACCGCAGCGGTAAAGACACCAAGTCCCGCGCATACGCAG	661	
Db	2329	CTCACCTCATGCTCCCAAGCAGCGGCAAAAGCACCAAGGTCCCGGTGCATATGCA	2388	
QY	662	CTCAGGGCTACAAGGTGTGTGCTCAACCCCTCCGTGCTGCAACAATGGGCTTTGGTG	721	
Db	2389	CTCAGGGCTACAAGGTGTGTGCTCAACCCCTCTGTTGCTGCAACACTGGGCTTTGGTG	2448	
QY	722	CTTACATGTCCAAGGCCCATGGATGTATCTTAACATCAGGACTGGGGTGAGGACAATTA	781	
Db	2449	CTTACATGTCCAAGGCTCATGGATCGATCTCTAAATCATCAGGACCGGGGTGAGAACTTA	2508	
QY	782	CTACTGGACCGCATCAGTATTCACCTAGCGGAAGTTCCTTCCGACCGGGGTGT	841	
Db	2509	CCACTGGACCGCCATCACTACTCCTCACTAAGGAGTTCCTTCCGACCGGGGTGT	2568	
QY	842	CAGGGGTGCTTATGACATAATAATTTGTGACAGGTGCCACTCCACGATGCAACATCA	901	
Db	2569	CGGGGGCGCTTATGACATAATAATTTGTGACAGGTGCCACTCCACGATGCAACATCA	2628	
QY	902	TCTTGGGCAATTGCACTGCTTGGCAAGCAGACCGCGGGGCGAGACTGACTGTGTC	961	

Db	2629	TCTTGGGCACTCGGCACACTGCTCTTGACCAAGCAGAGACTGCGGGGCGAGACTGGTTGTGC	2688	
QY	962	TGCGCACCGCTACCCCTCGGGCTCGGTCACTGTGCCCCCATCTTACATCGAGGAGTTG	1021	
Db	2689	TGCGCACCGCCACCCCTCGGGCTCGGTCACTGTGCCCCCATCTTACATCGAGGAGTTG	2748	
QY	1022	CTCTGTCCACTACTCGGAGAGATCCCTTTTATGGCAAGGCTATTCCCTTTGAAGCAATTA	1081	
Db	2749	CTCTGTCCACTACTCGGAGAGATCCCTTTTATGGCAAGGCTATTCCCTTTGAAGTATCA	2808	
QY	1082	AGGGGGGAGAGACATCTCATCTTCTGCCACTCAAGAAGAGTGCAGAGCTGCGCGCAA	1141	
Db	2809	AGGGGGGAGAGACATCTCATCTTCTGCCACTCAAGAAGAGTGCAGAGACTGCGCGCAA	2868	
QY	1142	AACTGGTCCGCTTGGCGCTCAATGCGGTGCTTACTACCGGCGCTTGATGTGCTCGTCA	1201	
Db	2869	AGCTGGTCCGCTTGGCGCTCAATGCGGTGCTTACTACCGGCGCTTGATGTGCTCGTCA	2928	
QY	1202	TCCCGACCAAGTGGTGTGCTGCTGGCAACTGAAGCCCTCATGACCGGCTTTTACCG	1261	
Db	2929	TCCCGACCAAGTGGTGTGCTGCTGGCAACTGAAGCCCTCATGACCGGCTTATACCG	2988	
QY	1262	GGGACTTCGATTCGTTGATAGACTGCAACACTGTGTGTCAACCAGACAGTCTGACGCC	1321	
Db	2989	GGGACTTCGATTCGTTGATAGACTGCAACACTGTGTGTCAACCAGACAGTCTGATTCAGCC	3048	
QY	1322	TTGACCCCTACCTTCAACATTTGACCAATCAAGCTTCCCGAGATGCTGCTCCGCTACTC	1381	
Db	3049	TTGACCCCTACCTTCAACATTTGACCAATCAAGCTTCCCGAGATGCTGCTCCCGCTACTC	3108	
QY	1382	AACTCGGGGTAGGACTGGCAGAGGAAAGCCAGGCACTTACAGATTTTGGCAGCCGGGG	1441	
Db	3109	AACTCGGGGTAGGACTGGCAGAGGAAAGCCAGGCACTTACAGATTTTGGCAGCCGGGG	3168	
QY	1442	AGCGCTCTTCTGGCAATTTGATCGTCTGTCTCTGTGAGTGTCTATGACGCGGTTGTG	1501	
Db	3169	AGCGCTCTTCTGGCAATTTGATCGTCTGTCTCTGTGAGTGTCTATGACGCGGTTGTG	3228	
QY	1502	CTTGTGTATGAGCTTACGCGCGGAGACCAAGTGTAGGCTACGAGCTACATGAAACACCC	1561	
Db	3229	CTTGTGTATGAGCTTACGCGCGGAGACCAAGTGTAGGCTACGAGCTACATGAAACACCC	3288	
QY	1562	CGGGACTTCCCGGTGTCACCAAGACCATTTGAAATTTTGGAGGGCGCTTTTACGGGTCTCA	1621	
Db	3289	CGGGCTTCCCGGTGTCACCAAGACCATTTGAAATTTTGGAGGGCGCTTTTACAGGCTCA	3348	
QY	1622	CCCAATAGACGCGCTTCTTATCCAGACAAGAGAGTGGGGAACCTTCCCTATC	1681	
Db	3349	CTCATATAGATGCCACTTCTTATCCAGACAAGAGAGTGGGGAACCTTCCCTTACC	3408	
QY	1682	TGGTAGCGTACCAAGCCACCGTGTGCGGTAGAGCTCAAGCCCTCCCGCTGTGGGACC	1741	
Db	3409	TGGTAGCGTACCAAGCCACCGTGTGCGGTAGAGCTCAAGCCCTCCCGCTGTGGGACC	3468	
QY	1742	AGATGTGGAAGTGTGATCCGTCTCAAGCCACCCCTCCATGGGCCAACACCTCTGCTAT	1801	
Db	3469	AGATGTGGAAGTGTGATCCGTCTCAAGCCACCCCTCCATGGGCCAACACCTCTGCTAT	3528	
QY	1802	ATAGACTGGCGCTGTCCAGAAATGAAGTCAACCTGACGCAACCGAGTCAAGATATATCA	1861	
Db	3529	ACAGACTGGCGCTGTTCAGAAATGAAGTCAACCTGACGCAACCGAGTCAAGATATATCA	3588	
QY	1862	TGACATGTATGTGCGCTGACCTGGAGGTGCTCAACAGTACCTGGGTGCTGTTGGCGGG	1921	
Db	3589	TGACATGTATGTGCGCGGACCTGGAGGTGCTCAACAGTACCTGGGTGCTGTTGGCGGG	3648	
QY	1922	TTCTGGCTGCTTTGGCGCGTATTGCTTATCCACAGGCTGCGTGTGCTATAGTGTAGGA	1981	
Db	3649	TCCTGGCTGCTTTGGCGCGTATTGCTTATCCACAGGCTGCGTGTGCTATAGTGTAGGA	3708	
QY	1982	TTGTCTTGTGCGGAAAGCGGCAATCATACCGCAGAGGAGTCTCTTACCGGGAGTTCG	2041	
Db	3709	TCGTCTTGTGCGGAAAGCGGCAATCATACCTGACAGGAAAGTCTCTCTACCGAGAGTTCG	3768	

QY 2042 ATGAATGGAAGTGTCT 2059  
Db 3769 ATGAGATGGAAGTGTCT 3786

RESULT 13  
I08294  
LOCUS 108294 9185 bp DNA linear PAT 02-DEC-1994  
DEFINITION Sequence 1 from Patent EP 0388232.  
ACCESSION I08294  
VERSION I08294.1 GI:588994  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 9185)  
AUTHORS Houghton, M., Choo, Q.-L. and Kuo, G.  
TITLE NANBV diagnostics and vaccines  
JOURNAL Patent: EP 0388232-A1 1 19-SEP-1990;  
FEATURES Location/Qualifiers  
source 1. 9185  
/organism="unknown"  
/mol\_type="unassigned DNA"

ORIGIN  
Query Match 86.7%; Score 1786; DB 6; Length 9185;  
Best Local Similarity 91.7%; Pred. No. 0;  
Matches 1888; Conservative 0; Mismatches 170; Indels 0; Gaps 0;

QY 2 TGGCGCTATACGGCTATGCCAGCAGACAAAGGGGCTTTGGGATGCATAATCACCA 61  
Db 3396 TGGCGCCCATCACGGGCTAGCCCCAGCAGACAAAGGGGCTTCTTAGGGTGCAATACCA 3455

QY 62 GCTTGACCGGCGCGGACAAAACACGAGTGAGGGTGAGGTTGAGCTGTCGTAACCTGCTG 121  
Db 3456 GCCTAACTGCGCGGACAAAACCAAGTGGAGGGTGAGGTTGTCACATGCTGCTG 3515

QY 122 CCAGACATTTCTTGGCAACTGATTAACGGGGTGTTGGAGTGTCTACCATGGAGCGG 181  
Db 3516 CCNAACTTCTTGGCAACTGATTAACGGGGTGTTGGAGTGTCTACCATGGAGCGG 3575

QY 182 GAACAAGGACCATTTGGTCACTTAAGGGTCTCTTATCCAGATGTACCAATATGGAGC 241  
Db 3576 GAACGAGGACCATTTGGTCACTTAAGGGTCTCTTATCCAGATGTACCAATATGGAGC 3635

QY 242 AAGACCTGTAGCTGGCGCGCTCCCAAGGTGCCCGCTCATTAACACATGACATCTGG 301  
Db 3636 AAGACCTGTGGGCTGGCGCGCTCCGCAAGGTAGCGGCTCATTTGACACCCCTGCACTTGG 3695

QY 302 GCTCTCGGACCTTTACCTGGTCAAGGACGACCGCATGTCATTCCTGTGGCGGACGG 361  
Db 3696 GCTCTCGGACCTTTACCTGGTCAAGGACGACCGCATGTCATTCCTGTGGCGGCGGG 3755

QY 362 GTGATGCGAGGGGAGCGCTCTTTGCGCCCGGCTATCTCTTAAAGGCTCCTCGG 421  
Db 3756 GTGATGCGAGGGGAGCGCTCTCTGCGCCCGGCTATCTCTTAAAGGCTCCTCGG 3815

QY 422 GAGGCGCTCTGTGTGCGCGGAGGACATGCCGTAGGCATATTCAGAGCGCGGATGCA 481  
Db 3816 GGGGTCCTGTGTGCGCGGAGGACATGCCGTAGGCATATTTAGGGCGCGGCTGTGCA 3875

QY 482 CCGGTGGAGTGGCTAAGGCGGTGAGCTTCATCCCGTAGAGCTTAGAGACAACCATGA 541  
Db 3876 CCGGTGGAGTGGCTAAGGCGGTGAGCTTCATCCCGTAGAGCTTAGAGACAACCATGA 3935

QY 542 GGTCCCGGCTGTCTCAGACAACTCTCTCCACAGCAGTGGCCCGAGAGCTACCAAGTGG 601  
Db 3936 GGTCCCGGCTGTCTCAGACAACTCTCTCCACAGCAGTGGCCCGAGAGCTTCAGGTGG 3995

QY 602 CCCACTGTGATGCTCCACCGGAGCGGTAAAGACCAAGAGTCCCGGCGCATACGAG 661  
Db 3996 CTCACCTCCATGCTCCACAGGAGCGGCAAAAGCAACAGGTCCCGGCTGCATATGCGAG 4055

QY 662 CTCAGGGCTACAAGGTGTGTGCTCAACCCCTCCGTTGCTCAACAATATGGCTTTGGTG 721  
Db 4056 CTCAGGGCTATAAGGTGTGTGCTCAACCCCTCCGTTGCTCAACAATATGGCTTTGGTG 4115

QY 722 CTTATATGTCCAAGGCCCATGGGATGATCTCTAAACATCAGGACTGGGGTGAGACAATTA 781  
Db 4116 CTTATATGTCCAAGGCTCATGGGATGATCTCTAAACATCAGGACTGGGGTGAGACAATTA 4175

QY 782 CTACTGGGACCGCCGATCAGTATTCACCTACGCAAGTTCCTTGGCGACGGCGGTGT 841  
Db 4176 CCACTGGGACCGCCCATCAGTATTCACCTACGCAAGTTCCTTGGCGACGGCGGTGT 4235

QY 842 CAGGGGCTGCTTATGACATAATAATTTGTGACAGTGGCCACTCCACGGATGCAACATCA 901  
Db 4236 CGGGGGGCGCTTATGACATAATAATTTGTGACAGTGGCCACTCCACGGATGCAACATCA 4295

QY 902 TCTTGGGCAATTTGGCACTGCTCTTACCAAGCAGAGACCGGGGGGAGACTGACTGTGC 961  
Db 4296 TCTTGGGCAATTTGGCACTGCTCTTACCAAGCAGAGACTCGGGGGGAGACTGACTGTGC 4355

QY 962 TCGCCACGCTACCCCTCGGGCTCGTCACTGTGCCCATCTTAACATCGAGGAGGTG 1021  
Db 4356 TCGCCACGCTACCCCTCGGGCTCGTCACTGTGCCCATCTTAACATCGAGGAGGTG 4415

QY 1022 CTCTGTCCACTACCGGAGAGATCCCTTTTATGGCAAGCTATTCCTCTTGAAGCAATTA 1081  
Db 4416 CTCTGTCCACTACCGGAGAGATCCCTTTTATGGCAAGCTATTCCTCTTGAAGTAATCA 4475

QY 1082 AGGGGGGAGACATCTCACTTCTGCTGCTCAAAAGAGAGTGGCGAGCTGCCCGCAA 1141  
Db 4476 AGGGGGGAGACATCTCACTTCTGCTCAAAAGAGAGTGGCGAGCTGCCCGCAA 4535

QY 1142 AACTGTGCGTGTGGCGTCAATGCGGTGCTTACTACCGCGCTTGTATGTGTCGTCA 1201  
Db 4536 AGTGTGCGCATTTGGGCTCAATGCGGTGCTTACTACCGCGCTTGTATGTGTCGTCA 4595

QY 1202 TCCGACCAAGTGTGTCGTGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCG 1261  
Db 4596 TCCGACCAAGTGTGTCGTGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCG 4655

QY 1262 GCGACTTCGATTCGCTGATAGACTGCAACAGTGTGTACCCAGACAGTGCATTCAGCC 1321  
Db 4656 GCGACTTCGATTCGCTGATAGACTGCAACAGTGTGTACCCAGACAGTGCATTCAGCC 4715

QY 1322 TTGACCTTACCTTCCATTGAGACAAATCAAGCTTCCCGAGATGCTCTCCCGTACTC 1381  
Db 4716 TTGACCTTACCTTCCATTGAGACAAATCAAGCTTCCCGAGATGCTCTCCCGTACTC 4775

QY 1382 AACGTGCGGGTAGGACTGGCAGAGGGAAGCCAGGCAATCTACAGATTTGTGGCACCGGG 1441  
Db 4776 AACGTGCGGGTAGGACTGGCAGAGGGAAGCCAGGCAATCTACAGATTTGTGGCACCGGG 4835

QY 1442 AGGCTCTTCTGCGATGTTGATCTGCTGTCTCTGCGAGTGTCTATGACGCGGTGTG 1501  
Db 4836 AGGCTCTTCTGCGATGTTGATCTGCTGTCTCTGCGAGTGTCTATGACGCGGTGTG 4895

QY 1502 CTTGGTATGAGCTTACGCGCGGAGACCAAGTGTAGGCTACGAGCATATCAAGACACCC 1561  
Db 4896 CTTGGTATGAGCTTACGCGCGGAGACCAAGTGTAGGCTACGAGCATATCAAGACACCC 4955

QY 1562 CGGGACTTCCCGTGTGCCAAGACCAATTTTGGAGGGGCTCTTTTACGGGTCTCA 1621  
Db 4956 CGGGACTTCCCGTGTGCCAAGACCAATTTTGGAGGGGCTCTTTTACGGGTCTCA 5015

QY 1622 CCCACATAGACGCCCATCTTCTATCCCAAGAGAGTGGGGGAAAACCTTCCCTATC 1681  
Db 5016 CTCATATAGATGCCCATCTTCTATCCCAAGAGAGTGGGGGAAAACCTTCCCTATC 5075

QY 1682 TGGTAGCTACCAAGCCAGCTGTGCGCTAGAGCTCAAGCCCTCCCGCTCGTGGGACC 1741  
Db 5076 TGGTAGCTACCAAGCCAGCTGTGCGCTAGAGCTCAAGCCCTCCCGCTCGTGGGACC 5135

```
QY 1742 AGATGTGGAAGTCTTGATCCGTCCTCAAGCCACCCCTCCATGGGCCCAACACCTCTGCTAT 1801
Db 5136 AGATGTGGAAGTCTTGATTCGGCTCAAGCCACCCCTCCATGGGCCCAACACCCCTCTAT 5195
QY 1802 ATAGACTGGGCGCTGTCAGAAATGAAGTCAACCTGACGACCCAGTCACCAAGATATATCA 1861
Db 5196 ACAGACTGGGCGCTGTCAGAAATGAAGTCAACCTGACGACCCAGTCACCAATATATCA 5255
QY 1862 TGACATGTATGTCGGCTGACCTGGAGGTGTCACGAGTACCTGGGTGCTGTTGGCGGG 1921
Db 5256 TGACATGTATGTCGGCGGACCTGGAGGTGTCACGAGTACCTGGGTGCTGTTGGCGGG 5315
QY 1922 TTCTGCTGCTTTGGCGGCTATGCTATCCACAGGCTGCGTGCATAGTAGTAGGA 1981
Db 5316 TCTGCTGCTTTGGCGGCTATGCTGTCACAGGCTGCGTGCATAGTAGTAGGA 5375
QY 1982 TTGCTTTGTCGGAAAGCGGCAATCATACCGGAGGAAGTCTCTACCGGGAGTTGCG 2041
Db 5376 TCGTCTTTGTCGGAAAGCGGCAATCATACCTGACAGGGAAGTCTCTACCGGAGATTGCG 5435
QY 2042 ATGAATGGAAGTGTCT 2059
Db 5436 ATGAGATGGAAGTGTCT 5453

RESULT 14
BD091382
LOCUS BD091382 9185 bp DNA linear PAT 27-AUG-2002
DEFINITION HCV cultivation method in eucaryotic cells.
ACCESSION BD091382
VERSION BD091382.1 GI:22636993
KEYWORDS JP 2001314192-A/3.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 9185)
Weiner,A.J., Steimer,K.S. and Houghton,M.
HCV cultivation method in eucaryotic cells
Patent: JP 2001314192-A 3 13-NOV-2001;
CHIRON CORP
OS Homo sapiens (human)
PN JP 2001314192-A/3
PD 13-NOV-2001
PF 15-MAR-2001 JP 2001075144
PR 25-AUG-1989 US 398667
PI AMY J WEINER,KATHELYN S STEIMER,MICHAEL HOUGHTON PC
C12N15/09,C12N5/10,C12N7/00//C12N7/00,C12R1:93,C12N15/00, PC
CC HCV cultivation method in eucaryotic cells
C12N5/00
FH Key
FT source
FT 1. .9185
Location/Qualifiers
/organism='Homo sapiens (human)'.
1. .9185
/organism='Homo sapiens'
/mol_type='genomic DNA'
/db_xref='taxon:9606'

FEATURES
source
Query Match 86.7%; Score 1786; DB 6; Length 9185;
Best Local Similarity 91.7%; Pred. No. 0;
Matches 1898; Conservative 0; Mismatches 170; Indels 0; Gaps 0;

ORIGIN
QY 2 TGGCGGCTATACGGCCTATGCGGACAGACAAAGGGCCCTTTGGGATGCATAATCACCA 61
Db 3396 TGGCGGCCATACGGCGTACGGCCAGCAGACAAAGGGCCCTCTAGGGTGCATAATCACCA 3455
QY 62 GCTTGACCGCGGACAAAAACAGGTGGAGGTGAGGTTCAGATCGTGCAACTGCTG 121
Db 3456 GCTTAAGTGGCGGACAAAAACAGGTGGAGGTGAGGTTCAGATCGTGCAACTGCTG 3515
QY 122 CCCAGACTTCTTGGCAACCTGCATTTAACGGGGTGTGTTGGACTGTCTACCATGGAGCGG 181
```

```
Db 3516 CCCAAACCTTCTCGCAACGTGCATCAATGGGGTGTGTGGACTGTCTACCAAGGGGCGG 3575
QY 182 GAACAAGGACCAATTCGTCAACCTAAGGGTCTGTGTTATCCAGATGTACACCAATGTGGACC 241
Db 3576 GAACGAGGACCAATTCGTCAACCTAAGGGTCTGTGTTATCCAGATGTATACCAATGTAGACC 3635
QY 242 AAGACTCTGTAGGCTGGCGGCTCCCAAGAGTGGCCGCTCATTAACAACATGACTGTGGC 301
Db 3636 AAGACTCTTGTGGGCTGGCGGCTCCCAAGAGTGGCCGCTCATTAACAACATGACTGTGGC 3695
QY 302 GCTCTCTCGAGACCTTTACCTGTGTACAGAGCACCGCATGTCTATTCTGTGCCGCCGACGCG 361
Db 3696 GCTCTCTCGAGACCTTTACCTGTGTACAGAGCACCGCATGTCTATTCTGTGCCGCCGCGGG 3755
QY 362 GTGATGGCAGGGGCGAGCTGCTTTTCCGCCCGGCTTATCTTTACTTTGAAAGGCTCTCTCGG 421
Db 3756 GTGATAGCAGGGGCGAGCTGCTGTGCGCCCGGCCCATTTCTTACTTTGAAAGGCTCTCTCGG 3815
QY 422 GAGGCCCTCTGCTGTGCGCCCGGAGACATGCCGTAGGACATATTCAGAGCCGCGGTATGCA 481
Db 3816 GGGGTCCGCTGTGTGCGCCCGGCGGACGCGCGTGGGCATATTTAGGGCCGCGGTGTGCA 3875
QY 482 CCCGTGGAGTGGCTAAGCGGTGGACTTTTCATCCCGTAGAGAGCTTAGAGACAAACCATGA 541
Db 3876 CCCGTGGAGTGGCTAAGCGGTGGACTTTTATCCCTGTGGAGAACCTAGAGACAAACATGA 3935
QY 542 GGTCCCCGGTGTCTCAGACAACTCCTCCCCACAGAGTGTCCCCAGAGCTACCAAGTGG 601
Db 3936 GGTCCCCGGTGTCTCAGCGGTAACCTCCTCTCCACAGTAGTGTGCCCGCAGAGCTTCCAGGTGG 3995
QY 602 CCCACTGCATGCTCCACCGGAGGGGTAGAGCACCAAGTCCCGGCCGATACGAC 661
Db 3996 CTCACCTCATGCTCTCCACAGCAGCGGGAAGCACCAGGTCCCGGCTGCATATGACAG 4055
QY 662 CTCAGGGCTACAAGGTGTGTGTCTCAACCCCTCCGTGTGCTGCAACAATGGGCTTTGGTG 721
Db 4056 CTCAGGGCTATAAGTGTGTGTCTCAACCCCTCTGTGCTGCAACAATGGGCTTTGGTG 4115
QY 722 CTTAATGTCCAAGGCCCATGGGATGTATCTTAAATCATCAGAGCTGGGTGAGGACAAATTA 781
Db 4116 CTTAATGTCCAAGGGCTCATGGATGTATCTTAAATCATCAGAGCTGGGTGAGGACAAATTA 4175
QY 782 CTACTGGCAGCCCGATCAGTATTCACCTACGGCAAGTTCCTTCGCGCGCGGTGT 841
Db 4176 CCACTGGCAGCCCGATCAGTATTCACCTACGGCAAGTTCCTTCGCGCGCGGTGT 4235
QY 842 CAGGGGTGCTTTATGACATAATAATTTGTGACAGTGTCCACTCCACGGATGCAACATCCA 901
Db 4236 CGGGGGCGCTTATGACATAATAATTTGTGACAGTGTCCACTCCACGGATGCAACATCCA 4295
QY 902 TCTTGGGCATTTGGGCTGCTGTTGACCAAGCAGACCGCGGGGCGAGACTGTGCTGTGC 961
Db 4296 TCTTGGGCATTCGGCACTGCTGTTGACCAAGCAGAGACTTCGCGGGGCGAGACTGTGCTGTGC 4355
QY 962 TCGCCACCGTACCCCTCCGGGCTCCGTCACCTGTCGCCCATCTCAACATCGAGGAGTTG 1021
Db 4356 TCGCCACCGGCACTCCGCTCCGGGCTCCGTCACCTGTCGCCCATCTCAACATCGAGGAGTTG 4415
QY 1022 CTCTGTCCACTACCGGAGAGATCCCTTTTATGGCAAGGCTATTCCTTCTTGAAGCAATTA 1081
Db 4416 CTCTGTCCACCGGAGAGATCCCTTTTATGGCAAGGCTATTCCTTCTTGAAGTAATCA 4475
QY 1082 AGGGGGGAGACATCTCTCTGTCCTCACTCAAGAAGAGTGCAGAGCTCGCGGCA 1141
Db 4476 AGGGGGGAGACATCTCTCTCTGTCCTCACTCAAGAAGAGTGCAGAGCTCGCGGCA 4535
QY 1142 AACTGCTCGTGGGCGTCAATGCGGTGCTTACTACCGGCTCTGATGTGCTCGTCA 1201
Db 4536 AGCTGCTCGCATTTGGGCACTCAATGCGGTGCTTACTACCGGCTCTGATGTGCTCGTCA 4595
QY 1202 TCCCGACAGTGGTGAAGTGTGCGTGTGGCAACTGACGCCCTCATGACCGGCTTTACCG 1261
```



Db 4596 TC0GACACAGCGCGGATGTTGTCGTGTCGCAACCGATGCCCTCATGACCGGCTATACCG 4655  
Qy 1262 GCAGCTTCGATTCGGGTATAGACTGCAACACGCTGTGTACCCAGACAGTCGACTTCAGCGC 1321  
Db 4656 GCGACTTCGACTCGGTATAGACTGCAATACGTTGTGTACCCAGACAGTCGATTCAGCGC 4715  
Qy 1322 TTGACCTTACCTTCACCATTTGAGACAAATACGCTTCCCGAGATGCTGTCTCCGTCAC 1381  
Db 4716 TTGACCTTACCTTCACCATTTGAGACAAATACGCTTCCCGAGATGCTGTCTCCGTCAC 4775  
Qy 1382 AAGTCTGGGTAGGACTTGGCAGAGGAAGCCAGCATCTACAGATTTGTGGACACCGGGG 1441  
Db 4776 AAGCTGGGCGAGACTTGGCAGGGAAGCCAGCATCTACAGATTTGTGGACACCGGGG 4835  
Qy 1442 AGCGTCTCTCTGTCATGTTTGAATCTGCTGTCTCTGTCAGAGTCTATGACGCGGGTTGTG 1501  
Db 4836 AGCGCCCTCCGCGATGTCGACTCGTCCGTCTCTGTGAGTCTATGACGAGGCTGTG 4895  
Qy 1502 CTTGGTATGAGCTTACGCGCGGAGACCAACAGTATAGGCTACGAGCATCATGAACACCC 1561  
Db 4896 CTTGGTATGAGCTTACGCGCGGAGACTACAGTATAGGCTACGAGCGTACATGAACACCC 4955  
Qy 1562 CGGAGCTTCCCGTGTCCAGACCATCTTGAATTTTGGAGGGGCTCTTACGGTCTCA 1621  
Db 4956 CGGGGCTTCCCGTGTCCAGACCATCTTGAATTTTGGAGGGGCTCTTACAGGCTCA 5015  
Qy 1622 CCACATAGACGCCACTTCTATCCAGACAAAGCAGAGTGGGGAACCTTCCCTATC 1681  
Db 5016 CTATATAGTGCCTTCTATCCAGACAAAGCAGAGTGGGGAACCTTCTTACC 5075  
Qy 1682 TGTAGCGTACCAAGCCACGCTGCGCTAGAGCTCAAGCCCTTCCCGCTCGTGGGACC 1741  
Db 5076 TGTAGCGTACCAAGCCACGCTGCGCTAGAGCTCAAGCCCTTCCCGCTCGTGGGACC 5135  
Qy 1742 AGATGTGGAAGTCTGTGATCCGCTCAAGCCACCTTCCAGAGGCTCTGCTAT 1801  
Db 5136 AGATGTGGAAGTCTGTGATCCGCTCAAGCCACCTTCCAGAGGCTCTGCTAT 5195  
Qy 1802 ATAGACTGGGCGCTGTCCAGAAATGAAGTCAACCTGACGACACCCAGTCAACAGTATATCA 1861  
Db 5196 ACAGACTGGGCGCTGTTCAGATGAATCAACCTGACGACACCCAGTCAACNAATATCA 5255  
Qy 1862 TGACATGTATGTCGGGTGACCTGGAGGTGTCAAGAGTACCTGGGTGCTGTGGCGCG 1921  
Db 5256 TGACATGTATGTCGGCGACCTGGAGGTGTCAAGAGTACCTGGGTGCTGTGGCGCG 5315  
Qy 1922 TTCTGCTGCTTTGGCGCGTATTGCTATCCAGAGCTGCGGTGCTCATAGTAGTAGGA 1981  
Db 5316 TCCTGCTGCTTTGGCGCGTATTGCTGCTCAAGGCTGCGGTGCTCATAGTAGTAGGA 5375  
Qy 1982 TTGCTTTGTCGGAAAGCGGCAATCATACCCGACAGGGAAGTCTCTACCGGAGTTGG 2041  
Db 5376 TCGTCTTTGTCGGGAAGCGGCAATCATACCTGACAGGGAAGTCTCTACCGGAGTTGG 5435  
Qy 2042 ATGAAATGGAAGAGTGTCT 2059  
Db 5436 ATGAGATGGAAGAGTGTCT 5453

RESULT 15  
AR166930  
LOCUS Sequence 1 from patent US 6284249.  
DEFINITION AR166930  
ACCESSION AR166930  
VERSION AR166930.1 GI:16243325  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unclassified.  
REFERENCE 1 (bases 1 to 9379)  
AUTHORS Barban, V.  
TITLE Fusion polypeptide having the C protein and E1 protein of hepatitis C virus

JOURNAL Patent: US 6284249-A 1 04-SEP-2001;  
FEATURES Location/Qualifiers  
source 1..9379  
/organism="unknown"  
/mol\_type="unassigned DNA"  
ORIGIN  
Query Match 86.7%; Score 1786; DB 6; Length 9379;  
Best Local Similarity 91.7%; Pred. No. 0;  
Matches 1888; Conservative 0; Mismatches 170; Indels 0; Gaps 0;  
Qy 2 TGGCGCTTATCAGCGCTATGCCAGCAGACAAAGGGGCTTTTGGATGTCATAATCACCA 61  
Db 3396 TGGCGCTTATCAGCGCTATGCCAGCAGACAAAGGGGCTTCTTGGGTGTCATAATCACCA 3455  
Qy 62 GCTTCACCGCGCGGACAAAACACAGGTGGAGGTGAGGTTCCAGATCGTGTCAACTGCTG 121  
Db 3456 GCTTAACTGGCGCGGACAAAACACAGGTGGAGGTGAGGTTCCAGATCGTGTCAACTGCTG 3515  
Qy 122 CCCAGACTTTCTTGGCAACTGTCATTAAAGGGGTGTGTGGAGTGTCTACCATGGAGCGG 181  
Db 3516 CCCAAACCTTCTGGCAAGCTGCATCAATGGGGTGTGTGGAGTGTCTACCATGGAGCGG 3575  
Qy 182 GAAAGAGACATTCGGTCACTTAAGGTCTCTGTTATCCAGATGTACCAATGTGGACC 241  
Db 3576 GAAAGAGACATTCGGTCACTTAAGGTCTCTGTTATCCAGATGTACCAATGTGGACC 3635  
Qy 242 AAGACTCGTAGGCTGGCGCGCTCCCAAGGTGCCGCTCATTAACACCATGCACCTTGGC 301  
Db 3636 AAGACTCGTAGGCTGGCGCGCTCCCAAGGTGCCGCTCATTAACACCATGCACCTTGGC 3695  
Qy 302 GCTCTCGGACCTTTACCTGTGTCAGAGGACAGCCGATGTCTATCTGTGCGCCGACGGG 361  
Db 3696 GCTCTCGGACCTTTACCTGTGTCAGAGGACAGCCGATGTCTATCTGTGCGCCGACGGG 3755  
Qy 362 GTGATGAGAGGGGACGCTGCTTTGCGCCCGGCTATCTCTTACTTTGAAAGGCTCTCGG 421  
Db 3756 GTGATGAGAGGGGACGCTGCTTTGCGCCCGGCTATCTCTTACTTTGAAAGGCTCTCGG 3815  
Qy 422 GAGGCGCTCTGCTGTGCGCGCAGGACATGCCGTAGGATATTTACAGAGCGCGGTATGCA 481  
Db 3816 GAGGCGCTCTGCTGTGCGCGCAGGACATGCCGTAGGATATTTACAGAGCGCGGTATGCA 3875  
Qy 482 CCGCTGGAGTGTGCTAAGCGGTGGACTTTCATCCCGTAGAGAGCTTAGAGACAACCATGA 541  
Db 3876 CCGCTGGAGTGTGCTAAGCGGTGGACTTTCATCCCGTAGAGAGCTTAGAGACAACCATGA 3935  
Qy 542 GGTCCCGGTGTCTCAGACAACTCTCTCCCAACAGAGTGTCCCGAGAGCTACCAAGTGG 601  
Db 3936 GGTCCCGGTGTCTCAGACAACTCTCTCTCCCAACAGAGTGTCCCGAGAGCTACCAAGTGG 3995  
Qy 602 CCCACTGCATGCTCCCAACCGGCTAAGAGCAGCAGGTCAGAGTCCCGCGCATACGCAG 661  
Db 3996 CTCACCTCCATGCTCCCAACCGGCTAAGAGCAGCAGGTCAGAGTCCCGCGCATACGCAG 4055  
Qy 662 CTCAGGCTTACAAGGTGTGCTGTCTCAACCCCTCCGTTGCTGCAACAATGGGCTTTGGTG 721  
Db 4056 CTCAGGCTTACAAGGTGTGCTGTCTCAACCCCTCCGTTGCTGCAACAATGGGCTTTGGTG 4115  
Qy 722 CTTACATGTCCAAGGCCCATGGATGTGATCTTAACATCAGGACTGGGTGAGACAATTA 781  
Db 4116 CTTACATGTCCAAGGCCCATGGATGTGATCTTAACATCAGGACTGGGTGAGACAATTA 4175  
Qy 782 CTACTGGAGCCCGGATCAGGATTTCCACTAGCGAAGTTCCTTCCGACGCGGGGTGT 841  
Db 4176 CCACTGGAGCCCGGATCAGGATTTCCACTAGCGAAGTTCCTTCCGACGCGGGGTGT 4235  
Qy 842 CAGGGGTGCTTATGACATAATAATTTGTGACAGTGCACCTCCACGAGTGAACATCCA 901  
Db 4236 CAGGGGTGCTTATGACATAATAATTTGTGACAGTGCACCTCCACGAGTGAACATCCA 4295  
Qy 902 TCTTGGGCAATGGCACTGTCTCTTGACCAAGAGAGACCGCGGGGCGAGACTGCTGTGC 961



Db 4296 TCTTGGGCAATCGGCACTGCTTCCTTGACCAAGACAGAGACTCGGGGGGAGAGACTGGTTGTGC 4355  
 Qy 962 TCGCCACCGCTACCCCTCGGGGCTCGGTCACTGTGCCCATCTTAACATCGAGGAGGTG 1021  
 Db 4356 TCGCCACCGGCAACCCCTCGGGGCTCGGTCACTGTGCCCATCCCAACATCGAGGAGGTG 4415  
 Qy 1022 CTCTGTCCACTACCGGAGAGATCCCTTTTATGGCAAGCTATTCCCTTTGAGCAATTA 1081  
 Db 4416 CTCTGTCCACCAACCGGAGAGATCCCTTTTATGGCAAGCTATTCCCTTTGAGCAATTA 4475  
 Qy 1082 AGGGGGGGAGACATCTCATCTTCTGCACTCAAGAGAAAGTGCAGAGCTCGCGCAA 1141  
 Db 4476 AGGGGGGGAGACATCTCATCTTCTGCACTCAAGAGAAAGTGCAGAGCTCGCGCAA 4535  
 Qy 1142 AACTGTCTCGGTGGGCGTCAATGCGGTGGCTTACTACCGCGGCTTGTATGTCCGTCA 1201  
 Db 4536 AGCTGTCTCGCATTTGGGCAATGCGGTGGCTTACTACCGCGGCTTGTATGTCCGTCA 4595  
 Qy 1202 TCCGACCAAGTGGTGAAGTGTGCTGCTGCAACTGACGCCCTCATGACCGGCTTTACCG 1261  
 Db 4596 TCCGACCAAGCGGCGATGTGCTGCTGCAACCGATGCCCTCATGACCGGCTTATACCG 4655  
 Qy 1262 GCGACTTCGATTCGGGTAGACTGCAACACGCTGTCTACCCAGACAGTCTGACTTCAGCC 1321  
 Db 4656 GCGACTTCGACTCGGTGTAGACTGCAATACGCTGTCTACCCAGACAGTCTGACTTCAGCC 4715  
 Qy 1322 TTGACCCCTACCTTACATGTAGACAATCAACGCTTCCCGAGGATGTCTCCCGTACTC 1381  
 Db 4716 TTGACCCCTACCTTACATGTAGACAATCAACGCTTCCCGAGGATGTCTCCCGTACTC 4775  
 Qy 1382 AACGTCGGGTAGGACTGCGAGAGGAGCCAGGCACTACAGATTTGGGCACCGGGG 1441  
 Db 4776 AACGTCGGGCGAGACTGGGAGGAGGAGCCAGGCACTACAGATTTGGGCACCGGGG 4835  
 Qy 1442 AGCGTCTCTGTCGATGTTTGAATCTCTCTGCGAGTCTATGACGCGGTTGTG 1501  
 Db 4836 AGCGCCCTCCGCAATGTTGCACTCGTCCGCTCTGTGAGTCTATGACGAGGCTGTG 4895  
 Qy 1502 CTTGGTATGAGCTTACCGCGCGGAGACACAGTATAGGCTACGAGCATACATGACACCC 1561  
 Db 4896 CTTGGTATGAGCTACGCGCGCGGAGACTACAGTATAGGCTACGAGCGTACATGACACCC 4955  
 Qy 1562 CGGGAATTCCTCGGTGCGCAAGACATCTTGAATTTTGGAGGCGCTTTTACGGGTCTCA 1621  
 Db 4956 CGGGGCTTCCGCTGTGCGAGACATCTTGAATTTTGGAGGCGCTTTTACGGGCTCA 5015  
 Qy 1622 CCCACATAGACGCCCACTTCTATCCAGACAAAGCAGAGTGGGAAACCTTCCCTATC 1681  
 Db 5016 CTCATATAGATGCCACTTTCTATCCAGACAAAGCAGAGTGGGAGAACCTTCCCTTACC 5075  
 Qy 1682 TGGTAGGTTACCAAGCCCGTGTGCGCTAGAGCTCAAGCCCTCCCGCTCGTGGGACC 1741  
 Db 5076 TGGTAGGTTACCAAGCCCGTGTGCGCTAGGCTCAAGCCCTCCCGCTCGTGGGACC 5135  
 Qy 1742 AGATGTGGAAGTGTCTGATCCGCTCAAGCCCACTTCCATGGGCGCAACCTCTGCTAT 1801  
 Db 5136 AGATGTGGAAGTGTCTGATCCGCTCAAGCCCACTTCCATGGGCGCAACCTCTGCTAT 5195  
 Qy 1802 ATAGACTGGGCGGTGTCCAGATGAAGTCAACCTGACGCAACCCAGTCAACCAATATCA 1861  
 Db 5196 ACAGACTGGGCGGTGTTCAGATGAATCAACCTGACGCAACCCAGTCAACCAATATCA 5255  
 Qy 1862 TGACATGTATGTGCGGTGAGCTGGAGTGTGCTGAGTACCTGGGTGCTGTTGGCGGG 1921  
 Db 5256 TGACATGTATGTGCGGCGCACTGGAGTGTGCTGAGTACCTGGGTGCTGTTGGCGGG 5315  
 Qy 1922 TTTGCGGTGCTTTGGCGCGCTATTTGCTATCCACAGGCTCGGTGCTATAGTAGGTAGGA 1981  
 Db 5316 TCTGGGTGCTTTGGCGCGCTATTTGCTGTCAACAGGCTCGGTGCTATAGTAGGTAGGA 5375  
 Qy 1982 TTGTCTTGTCCGGAAGCGGCAATCATACCGACAGGGAAGTCTCTACCGGGAGTTG 2041  
 Db 5376 TCGTCTTGTCCGGAAGCGGCAATCATACCTGACAGGGAAGTCTCTACCGAGAGTTG 5435

Qy 2042 ATGAANTGGAAGAGTCT 2059  
 Db 5436 ATGAGATGGAAGAGTCT 5453

Search completed: February 27, 2005, 20:54:59  
 Job time : 8691 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 27, 2005, 15:10:47 ; Search time 1053 Seconds  
(without alignments)  
11586.502 Million cell updates/sec

Title: US-09-930-591-1

Perfect score: 2061

Sequence: 1 atggcgccatcacggccta.....atgaaatggaagagtgtga 2061

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

N\_Geneseq\_16Dec04:\*

1: Geneseqn1980s:\*

2: Geneseqn1990s:\*

3: Geneseqn2000s:\*

4: Geneseqn2001as:\*

5: Geneseqn2001bs:\*

6: Geneseqn2002as:\*

7: Geneseqn2002bs:\*

8: Geneseqn2003as:\*

9: Geneseqn2003bs:\*

10: Geneseqn2003cs:\*

11: Geneseqn2003ds:\*

12: Geneseqn2004as:\*

13: Geneseqn2004bs:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2061	100.0	2061	6 AAD34500	Aad34500 Hepatitis
2	2061	100.0	2061	6 AAD31767	Aad31767 Hepatitis
3	2061	100.0	2061	10 AAD60868	Aad60868 Hepatitis
4	2061	100.0	2061	12 ADG47658	Adg47658 HCV NS3/4
5	1787	86.7	6299	4 AAF93689	Aaf93689 HCV NS3/4
6	1786	86.7	7310	1 AAN92106	Aan92106 Combined
7	1786	86.7	7310	1 AAN90336	Aan90336 Composite
8	1786	86.7	7310	2 AAQ98221	Aaq98221 Hepatitis
9	1786	86.7	8316	3 AAA75296	Aaa75296 cDNA sequ
10	1786	86.7	9133	2 AAQ07656	Aaz07656 Nucleotid
11	1786	86.7	9185	2 AAQ10566	Aaq10566 Hepatitis
12	1786	86.7	9185	3 AAA75297	Aaa75297 Sense str
13	1786	86.7	9185	12 AAD35979	Aad35979 HCV cDNA
14	1786	86.7	9401	2 AAT12710	Aat12710 Hepatitis
15	1786	86.7	9401	2 AAT99981	Aat99981 HCV polyp
16	1786	86.7	9401	2 AAV09989	Aav09989 HCV polyp
17	1786	86.7	9401	6 AAD35043	Aad35043 Hepatitis
18	1786	86.7	9401	12 ADL23106	Adl23106 Hepatitis
19	1786	86.7	9401	13 ADR29358	Adr29358 Hepatitis
20	1784.4	86.6	6905	1 AAN92103	Aan92103 Combined

## ALIGNMENTS

### RESULT 1

AAD34500

ID AAD34500 standard; DNA; 2061 BP.

XX

AC AAD34500;

XX

DT 16-JUL-2002 (first entry)

XX

DE Hepatitis C virus NS3/4A protein encoding DNA.

XX

Hepatitis C virus; HCV; NS3/4A protein; therapy; HCV infection; vaccine;

KW virucide; gene; ds.

XX

OS Hepatitis C virus.

XX

PH Key Location/Qualifiers

FT CDS

FT /tag= a

FT /product= "HCV NS3/4A protein"

XX

PN WO200214362-A2.

XX

PD 21-FEB-2002.

XX

PF 15-AUG-2001; 2001WO-IB001774.

XX

PR 17-AUG-2000; 2000US-0225767P.

PR

29-AUG-2000; 2000US-0229175P.

PR

03-NOV-2000; 2000US-00705547.

XX

(TRIP-) TRIPEP AB.

XX

PI Sallberg M;

XX

WPI; 2002-339446/37.

DR

P-PSDB; AAE21837.

XX

Novel hepatitis C virus NS3/4A peptide useful for diagnosing presence or absence of hepatitis C virus in a subject and for preparing a medicament for treating hepatitis C virus infection.

PT

Claim 1; Page 64-65; 90pp; English.

XX

PS

CC The present invention relates to novel hepatitis C virus (HCV) NS3/4A  
 CC proteins and their corresponding polynucleotides. NS3/4A sequences are  
 CC useful for identifying the presence or absence of HCV in a subject. They are  
 CC useful for preparing a medicament used for treating or preventing HCV  
 CC infection. Sequences of the invention are also used as vaccines. The  
 CC present sequence is a DNA encoding HCV NS3/4A protein  
 XX  
 SQ Sequence 2061 BP; 427 A; 616 C; 571 G; 447 T; 0 U; 0 Other;

		Query Match	100.0%;	Score 2061;	DB 6;	Length 2061;		
		Best Local Similarity	100.0%;	Pred. No. 0;				
		Matches 2061;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;		
QY	1	ATGGCGCCCTATACAGCCCTATGCGCCAGCAGACAAAGGGGCTTTTGGGATGCATAATCAC	60					
DB	1	ATGGCGCCCTATACAGCCCTATGCGCCAGCAGACAAAGGGGCTTTTGGGATGCATAATCAC	60					
QY	61	AGCTTGACCGCGCGGACAAAACACAGGTGGAGGTGAGGTTCAGATCGTGTCAACTGCT	120					
DB	61	AGCTTGACCGCGCGGACAAAACACAGGTGGAGGTGAGGTTCAGATCGTGTCAACTGCT	120					
QY	121	GCCACAGCTTCTTGGAACCTGCAATTAACGGGGTGTGTGGACTGTCTACCATGGAGCC	180					
DB	121	GCCACAGCTTCTTGGAACCTGCAATTAACGGGGTGTGTGGACTGTCTACCATGGAGCC	180					
QY	181	GGAACAGGACCAATTCGCTCACTAAGGTCTCTTATCCAGATGTACACCAATGTGGAC	240					
DB	181	GGAACAGGACCAATTCGCTCACTAAGGTCTCTTATCCAGATGTACACCAATGTGGAC	240					
QY	241	CAAGACCTCGTAGGCTGGCCGCTCCCAAGGTGCGGCTCATTAACCATGCACTTGC	300					
DB	241	CAAGACCTCGTAGGCTGGCCGCTCCCAAGGTGCGGCTCATTAACCATGCACTTGC	300					
QY	301	GGCTCTCGGACCTTACCTGTGTCAGGACGAGCGCGATGTCAATTCCTGTGCGCGAGG	360					
DB	301	GGCTCTCGGACCTTACCTGTGTCAGGACGAGCGCGATGTCAATTCCTGTGCGCGAGG	360					
QY	361	GGTGATGGCAGGGGAGCTGCTTTCGCGCCGCTCTCTTACTTGAAGGCTCCTCG	420					
DB	361	GGTGATGGCAGGGGAGCTGCTTTCGCGCCGCTCTCTTACTTGAAGGCTCCTCG	420					
QY	421	GGAGGCGCTCTGTGTCGCGCGAGCATGCGTAGGCATATTCAGAGCGCGGTATGC	480					
DB	421	GGAGGCGCTCTGTGTCGCGCGAGCATGCGTAGGCATATTCAGAGCGCGGTATGC	480					
QY	481	ACCGTGGAGTGGCTAAGCGGTGACCTTCATCCCGTAGAGCTTAGAGCAACCATG	540					
DB	481	ACCGTGGAGTGGCTAAGCGGTGACCTTCATCCCGTAGAGCTTAGAGCAACCATG	540					
QY	541	AGGTCCCGGTGTTCTCAGACAACTCCTCCCAACAGCAGTGCCTCCAGAGTACCAGTG	600					
DB	541	AGGTCCCGGTGTTCTCAGACAACTCCTCCCAACAGCAGTGCCTCCAGAGTACCAGTG	600					
QY	601	GCCACCTGCATGCTCCCAACCGGACGCTAAGAGCACCAAGTCCCGCGCATACGCA	660					
DB	601	GCCACCTGCATGCTCCCAACCGGACGCTAAGAGCACCAAGTCCCGCGCATACGCA	660					
QY	661	GCTCAGGGTACAAGGTGCTGTGCTCAACCCCTCGTGTGCTGCAACATGGGCTTTGGT	720					
DB	661	GCTCAGGGTACAAGGTGCTGTGCTCAACCCCTCGTGTGCTGCAACATGGGCTTTGGT	720					
QY	721	GCTTACATGCTCAAGGCCCATGGGATTTGATCTTAACATCAGGACTGGGGTAGGCAATT	780					
DB	721	GCTTACATGCTCAAGGCCCATGGGATTTGATCTTAACATCAGGACTGGGGTAGGCAATT	780					
QY	781	ACTACTGGACCGCATACGATTTCCACTACGGCAAGTTCCTTGGCAGCGCGGTGT	840					
DB	781	ACTACTGGACCGCATACGATTTCCACTACGGCAAGTTCCTTGGCAGCGCGGTGT	840					
QY	841	TCAGGGGTGCTTATGACATAATAATTGTGACGAGTGCCACTCCACGGATGCAATCC	900					
DB	841	TCAGGGGTGCTTATGACATAATAATTGTGACGAGTGCCACTCCACGGATGCAATCC	900					

QY	901	ATCTTTGGGCAITGGCACCTGTCTTGAACAGCAGAGACCGCGGGGCGAGACTGACTGTG	960					
DB	901	ATCTTTGGGCAITGGCACCTGTCTTGAACAGCAGAGACCGCGGGGCGAGACTGACTGTG	960					
QY	961	CTCGCCACCGCTACCCCTCCGGGCTCCGCTCACTGTGCCCATCTTAACATCAGGAGGTT	1020					
DB	961	CTCGCCACCGCTACCCCTCCGGGCTCCGCTCACTGTGCCCATCTTAACATCAGGAGGTT	1020					
QY	1021	GCTCTGTCCACTACCGGAGAGATCCCTTTTATGCAAGGCTATTCCCTTGAAGCAATT	1080					
DB	1021	GCTCTGTCCACTACCGGAGAGATCCCTTTTATGCAAGGCTATTCCCTTGAAGCAATT	1080					
QY	1081	AAGGGGGGAGACATCTCATCTTCTGCCACTCAAAAGAGAGTGCAGAGCTCCCGCA	1140					
DB	1081	AAGGGGGGAGACATCTCATCTTCTGCCACTCAAAAGAGAGTGCAGAGCTCCCGCA	1140					
QY	1141	AAACTGGTGGCTTGGGCTCAATCCGCTGGCTTACTACCGGCTTGTATGTGTCCGTC	1200					
DB	1141	AAACTGGTGGCTTGGGCTCAATCCGCTGGCTTACTACCGGCTTGTATGTGTCCGTC	1200					
QY	1201	ATCCCGACAGTGGT	1260					
DB	1201	ATCCCGACAGTGGT	1260					
QY	1261	GGCGACTTTCGATTTCGCTGATAGACTGCAACACGTGTGTGTGTGTGTGTGTGTGTGT	1320					
DB	1261	GGCGACTTTCGATTTCGCTGATAGACTGCAACACGTGTGTGTGTGTGTGTGTGTGTGT	1320					
QY	1321	CTTGACCTTACCTTCCACCAATTTGAGCAATACGCTTCCCGAGGATGTGTCTCCGTA	1380					
DB	1321	CTTGACCTTACCTTCCACCAATTTGAGCAATACGCTTCCCGAGGATGTGTCTCCGTA	1380					
QY	1381	CAAGCTCGGGGTAGAGCTGGCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	1440					
DB	1381	CAAGCTCGGGGTAGAGCTGGCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	1440					
QY	1441	GAGCGCTCTTCTGGCATGT	1500					
DB	1441	GAGCGCTCTTCTGGCATGT	1500					
QY	1501	GCTTGGTATGAGCTTACCGCGCGGAGACCAAGTGTGTGTGTGTGTGTGTGTGTGTGT	1560					
DB	1501	GCTTGGTATGAGCTTACCGCGCGGAGACCAAGTGTGTGTGTGTGTGTGTGTGTGTGT	1560					
QY	1561	CGGGACCTTCCGCTGTGCGAGACCACTTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	1620					
DB	1561	CGGGACCTTCCGCTGTGCGAGACCACTTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	1620					
QY	1621	ACCCACATAGACGCGGCTTCTTATCCAGACAAAGAGAGTGGGGAACCTTCCCTAT	1680					
DB	1621	ACCCACATAGACGCGGCTTCTTATCCAGACAAAGAGAGTGGGGAACCTTCCCTAT	1680					
QY	1681	CTGTGAGCTTACCAAGCCACCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	1740					
DB	1681	CTGTGAGCTTACCAAGCCACCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	1740					
QY	1741	CAGATGTGGAGT	1800					
DB	1741	CAGATGTGGAGT	1800					
QY	1801	TATAGACTGGGCGCTGTCCAGAAATGAAGTCAACCTGTGTGTGTGTGTGTGTGTGTGT	1860					
DB	1801	TATAGACTGGGCGCTGTCCAGAAATGAAGTCAACCTGTGTGTGTGTGTGTGTGTGTGT	1860					
QY	1861	ATGACATGTATGTGGGCTGACCTGGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	1920					
DB	1861	ATGACATGTATGTGGGCTGACCTGGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	1920					
QY	1921	GTTCTGGCTGT	1980					
DB	1921	GTTCTGGCTGT	1980					
QY	1981	ATTGTCTTGTCCGGAAGCCGCGCAATCATATCCCGACAGGAGTCTCTTACCGGAGTTC	2040					

Db 1981 ATTGTTCTTGTCCGGAAGCCGCAATCATACCGACAGGAAGTCTCTACCGGGAGTTC 2040  
QY 2041 GATGAATGGAAGAGTGCTGA 2061  
Db 2041 GATGAATGGAAGAGTGCTGA 2061  
RESULT 2  
AAD31767  
ID AAD31767 standard; DNA; 2061 BP.  
AC AAD31767;  
XX  
XX 18-JUN-2002 (first entry)  
XX Hepatitis C virus (HCV) NS3/4A DNA coding region.  
DE  
XX  
XX Hepatitis C virus; HCV infection; virucide; fungicide; antibacterial;  
KW cytotatic; immunostimulant; vaccine; ribavirin; immune response; cancer;  
KW ds.  
XX  
OS Hepatitis C virus.  
XX  
XX Key Location/Qualifiers  
FH CDS 1..2061  
FT /\*tag= a  
FT /product= "HCV NS3/4A protein"  
XX  
PN WO200213855-A2.  
XX  
XX 21-FEB-2002.  
XX  
XX 15-AUG-2001; 2001WO-1B001808.  
XX  
XX 17-AUG-2000; 2000US-0225767P.  
PR 29-AUG-2000; 2000US-0229175P.  
PR 03-NOV-2000; 2000US-00705547.  
XX  
XX (TRIP-) TRIPEP AB.  
XX  
XX Sallberg M, Hultgren C;  
PI  
XX WPI; 2002-241837/29.  
DR P-PSDB; AAE19900.  
XX  
XX Vaccine compositions for treating and preventing disease, preferably  
PT hepatitis C virus infection, comprises ribavirin and antigen that has  
PT epitope present in hepatitis C virus.  
XX  
XX Claim 1; Page 94-95; 120pp; English.  
XX  
XX The invention relates to a composition comprising ribavirin and an  
CC antigen preferably non structural 3 protein (NS3)/4A fragment of  
CC hepatitis C virus (HCV) genome or a peptide or nucleic acid of HCV  
CC sequence. The composition is useful for enhancing an immune response to a  
CC hepatitis C antigen in humans, domestic, sport or pet species and as  
CC vaccines for treating and preventing HCV infections. The composition is  
CC also useful for treating viral, bacterial, fungal diseases and cancer.  
CC The present sequence is HCV NS3/4A DNA coding region  
XX  
SQ Sequence 2061 BP; 427 A; 616 C; 571 G; 447 T; 0 U; 0 Other;  
Query Match 100.0%; Score 2061; DB 6; Length 2061;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 2061; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 ATGGCGCCTATCAGCGCCTATGCCAGACAGCAAGGGGCTTTTGGGATGCATATCACC 60  
Db 1 ATGGCGCCTATCAGCGCCTATGCCAGACAGCAAGGGGCTTTTGGGATGCATATCACC 60  
QY 61 AGCTTGACCGCGCGGACAAAACCAAGGTGAGGGTTCAGATCGTGCAACTGCT 120

Db 61 AGCTTGACCGCGCGGACAAAACCAAGGTGAGGGTTCAGATCGTGCAACTGCT 120  
QY 121 GCCCAGACTTTTCTTGGCAACCTGCAATTAACCGGGGTGTGTGGACTGTCTTACCATGAGCC 180  
Db 121 GCCCAGACTTTTCTTGGCAACCTGCAATTAACCGGGGTGTGTGGACTGTCTTACCATGAGCC 180  
QY 181 GGAACAAGGACCAATTGCGTCACTTAAGGGTCTCTGTATCCAGATGTACCAATGTGGAC 240  
Db 181 GGAACAAGGACCAATTGCGTCACTTAAGGGTCTCTGTATCCAGATGTACCAATGTGGAC 240  
QY 241 CAAGACCTCGTAGGCTGCGCCGCTCCCAAGGTGCGCGCTCAITTAACCATGCACTTGC 300  
Db 241 CAAGACCTCGTAGGCTGCGCCGCTCCCAAGGTGCGCGCTCAITTAACCATGCACTTGC 300  
QY 301 GGCTCTCGGACCTTTTACCTGTCACGAGGACGCGGATGTCAITTCCTGTGCGCGGACCG 360  
Db 301 GGCTCTCGGACCTTTTACCTGTCACGAGGACGCGGATGTCAITTCCTGTGCGCGGACCG 360  
QY 361 GGTGATGAGGAGGCGGAGCTGTTTGGCCCGCGCTATCTTACTTTGAAAGGCTCTCG 420  
Db 361 GGTGATGAGGAGGCGGAGCTGTTTGGCCCGCGCTATCTTACTTTGAAAGGCTCTCG 420  
QY 421 GGAGGCGCTCTGCTGTGCGCGCAGACATGCGGTAGGCATATTGAGAGCCGCGGTATGC 480  
Db 421 GGAGGCGCTCTGCTGTGCGCGCAGACATGCGGTAGGCATATTGAGAGCCGCGGTATGC 480  
QY 481 ACCCGTGGAGTGGCTAAGCGGTGAGACTTTCATCCCGTAGAGAGCTTAGAGACAACATG 540  
Db 481 ACCCGTGGAGTGGCTAAGCGGTGAGACTTTCATCCCGTAGAGAGCTTAGAGACAACATG 540  
QY 541 AGGTCCCGGTGTTCTCAGACAACCTCTCCCAAGGTGCGCGGAGCTTACCAAGTG 600  
Db 541 AGGTCCCGGTGTTCTCAGACAACCTCTCCCAAGGTGCGCGGAGCTTACCAAGTG 600  
QY 601 GCCCAGCTGCACTGCTCCAGCGGAGCGGTAGAGACCAAGGTCCCGCGCATACGCA 660  
Db 601 GCCCAGCTGCACTGCTCCAGCGGAGCGGTAGAGACCAAGGTCCCGCGCATACGCA 660  
QY 661 GCTCAGGGCTACAAGGTGCTGCTCAACCCCTCGTGTGCTGCAACATGGGCTTTGGT 720  
Db 661 GCTCAGGGCTACAAGGTGCTGCTCAACCCCTCGTGTGCTGCAACATGGGCTTTGGT 720  
QY 721 GCTTACATGTCOAAGGCCCATGGGATGTATCTCAATCAGAGCTGGGGTGAGGACAAT 780  
Db 721 GCTTACATGTCOAAGGCCCATGGGATGTATCTCAATCAGAGCTGGGGTGAGGACAAT 780  
QY 781 ACTACTGCGAGCCCATCAGTATTTCCACCTAGCGCAAGTTCCTTGCAGCGCGGTCT 840  
Db 781 ACTACTGCGAGCCCATCAGTATTTCCACCTAGCGCAAGTTCCTTGCAGCGCGGTCT 840  
QY 841 TCAGGGGGTGTCTTATGACATAATAATTTGTGACGAGTGCCACTCCACGGATGCAACATCC 900  
Db 841 TCAGGGGGTGTCTTATGACATAATAATTTGTGACGAGTGCCACTCCACGGATGCAACATCC 900  
QY 901 ATCTTTGGGCAATGGCACTGTCTTCCCAAGCAGAGACGCGGGGGCGAGACTGACTGTG 960  
Db 901 ATCTTTGGGCAATGGCACTGTCTTCCCAAGCAGAGACGCGGGGGCGAGACTGACTGTG 960  
QY 961 CTCGCCACCGCTACCCCTCCGGGCTCCGTCACTGTGCCCCCATCTTAAACATCGAGAGGTT 1020  
Db 961 CTCGCCACCGCTACCCCTCCGGGCTCCGTCACTGTGCCCCCATCTTAAACATCGAGAGGTT 1020  
QY 1021 GCTCTGTCCACTACCGGAGAGATCCCTTTTATGCAAGGCTATTTCCTTCAAGCAAT 1080  
Db 1021 GCTCTGTCCACTACCGGAGAGATCCCTTTTATGCAAGGCTATTTCCTTCAAGCAAT 1080  
QY 1081 AAGGGGGGAGACATCTCATCTTCTGCGCACTCAAGAGAAAGTGGAGCGAGCTGCGCGCA 1140  
Db 1081 AAGGGGGGAGACATCTCATCTTCTGCGCACTCAAGAGAAAGTGGAGCGAGCTGCGCGCA 1140  
QY 1141 AAATGCTGCGGCTGGGGCTCAATGCGGTGGCTTACTACCGCGGCTTCACTGTGCTCCGTC 1200  
Db 1141 AAATGCTGCGGCTGGGGCTCAATGCGGTGGCTTACTACCGCGGCTTCACTGTGCTCCGTC 1200

```
QY 1201 ATCCGACAGTGTGACGTTGTGTCGTGGCAACTGACGCGCTCATGACCGGCTTTACC 1260
Db 1201 ATCCGACAGTGTGACGTTGTGTCGTGGCAACTGACGCGCTCATGACCGGCTTTACC 1260
QY 1261 GCGGACTTCGATTTCGGTGATAGACTGCAACAGTGTGTGTCACCCAGACAGTGTGACG 1320
Db 1261 GCGGACTTCGATTTCGGTGATAGACTGCAACAGTGTGTGTCACCCAGACAGTGTGACG 1320
QY 1321 CTTGACCTTACCTTACCAATGAGACAATCAGCTTCCCGAGAGTGTCTCTCCGTAAT 1380
Db 1321 CTTGACCTTACCTTACCAATGAGACAATCAGCTTCCCGAGAGTGTCTCTCCGTAAT 1380
QY 1381 CAACGTCGCGGTAGGACTGCGAGAGGAAAGCAGGCACTCTACAGATTTGTGCGACCGGG 1440
Db 1381 CAACGTCGCGGTAGGACTGCGAGAGGAAAGCAGGCACTCTACAGATTTGTGCGACCGGG 1440
QY 1441 GAGCGTCTTCTGGCATGTTTGAATCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1500
Db 1441 GAGCGTCTTCTGGCATGTTTGAATCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1500
QY 1501 GCTTGTATGAGTTAGCGCGCGCGAGACACAGTTAGGCTAGGACATACATGAACACC 1560
Db 1501 GCTTGTATGAGTTAGCGCGCGCGAGACACAGTTAGGCTAGGACATACATGAACACC 1560
QY 1561 CCGGACTTCCGCTGTCGCAAGACCATCTTGAATTTTGGGAGGCGTCTTTACGGGTCTC 1620
Db 1561 CCGGACTTCCGCTGTCGCAAGACCATCTTGAATTTTGGGAGGCGTCTTTACGGGTCTC 1620
QY 1621 ACCACATAGAGCCCACTTCTATCCAGACAAAGCAGAGTGGGAAACCTTCCCTAT 1680
Db 1621 ACCACATAGAGCCCACTTCTATCCAGACAAAGCAGAGTGGGAAACCTTCCCTAT 1680
QY 1681 CTGTTAGCTACCAAGCCACGTTGCGGTAGAGCTCAAGCCCTTCCCGCTCGTGGGAC 1740
Db 1681 CTGTTAGCTACCAAGCCACGTTGCGGTAGAGCTCAAGCCCTTCCCGCTCGTGGGAC 1740
QY 1741 CAGATGTGAAGTGTGATGCTGCTCAAGCCACCTCCATGGGCGCAACCTCTGCTA 1800
Db 1741 CAGATGTGAAGTGTGATGCTGCTCAAGCCACCTCCATGGGCGCAACCTCTGCTA 1800
QY 1801 TATAGCTGGGCGCTGTCCAGAAATGAAGTCACTTCAAGCAGCCAGTCAACCAAGTATATC 1860
Db 1801 TATAGCTGGGCGCTGTCCAGAAATGAAGTCACTTCAAGCAGCCAGTCAACCAAGTATATC 1860
QY 1861 ATGACATGTATGTGCGCTGACCTGAGGTGCTGACAGTACCTGGTGTCTGTGGCGGC 1920
Db 1861 ATGACATGTATGTGCGCTGACCTGAGGTGCTGACAGTACCTGGTGTCTGTGGCGGC 1920
QY 1921 GTTCTGGCTGCTTTGGCGCGTATTGCTTATCCAGAGTGGTGTGCTATAGTAGG 1980
Db 1921 GTTCTGGCTGCTTTGGCGCGTATTGCTTATCCAGAGTGGTGTGCTATAGTAGG 1980
QY 1981 ATTGTCTTCTCCGAAAGCCGCAATATATACCCGACAGGGAAGTCTCTTACCGGAGTTTC 2040
Db 1981 ATTGTCTTCTCCGAAAGCCGCAATATATACCCGACAGGGAAGTCTCTTACCGGAGTTTC 2040
QY 2041 GATGAATGGAAGTGTCTGA 2061
Db 2041 GATGAATGGAAGTGTCTGA 2061
```

## RESULT 3

```
AAD60868
ID AAD60868 standard; DNA; 2061 BP.
XX
AC AAD60868;
XX
DT 15-JAN-2004 (first entry)
XX
DE Hepatitis C virus NS3/4A DNA.
XX
KW Ribavirin; vaccine; immune response; infection; therapy; immunostimulant;
```

```
virucide; ds.
XX
OS Hepatitis C virus.
XX
FH Key Location/Qualifiers
FT CDS 1..2061
FT /tag= a
FT /product= "Hepatitis C virus protein"
XX
PN US2002136740-A1.
XX
PD 26-SEP-2002.
XX
PF 15-AUG-2001; 2001US-00929955.
XX
PR 17-AUG-2000; 2000US-0225767P.
PR 29-AUG-2000; 2000US-0229175P.
XX
PA (SALL/) SALLBERG M.
PA (HULT/) HULTGREN C.
XX
PI Sallberg M, Hultgren C;
XX
DR WPI; 2003-764978/72.
DR P-PSDB; ABW00351.
XX
PT Vaccine compositions for treating and preventing disease, preferably
PT hepatitis C virus infection, comprises ribavirin and antigen that has
PT epitope present in hepatitis C virus.
XX
PS Claim 1; Page 60-61; Opp; English.
XX
CC The invention relates to a composition comprising ribavirin and an
CC antigen, where the antigen is derived from a hepatitis virus. The vaccine
CC is useful in enhancing the immune response to a hepatitis C antigen where
CC the composition is delivered to an animal identified as requiring an
CC enhanced immune response. The vaccine is useful in the treatment and
CC prevention of hepatitis C infection. The present sequence is Hepatitis C
CC virus NS3/4A DNA
XX
SQ Sequence 2061 BP; 427 A; 616 C; 571 G; 447 T; 0 U; 0 Other;
Query Match 100.0%; Score 2061; DB 10; Length 2061;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2061; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGGCGCTATCAGCGCTATGCCAGCAGACAAGGGGCTTTTGGGATGCATAATCACC 60
Db 1 ATGGCGCTATCAGCGCTATGCCAGCAGACAAGGGGCTTTTGGGATGCATAATCACC 60
QY 61 AGCTTGACCGCGCGGACAAAACAGGTGGAGGTGAGGTTTCAGATCGTGTCAACTGCT 120
Db 61 AGCTTGACCGCGCGGACAAAACAGGTGGAGGTGAGGTTTCAGATCGTGTCAACTGCT 120
QY 121 GCCCAGACTTCTTGGCAACCTGCAATTAACCGGGTGTGTGGACTGTCTTACCATGGAGCC 180
Db 121 GCCCAGACTTCTTGGCAACCTGCAATTAACCGGGTGTGTGGACTGTCTTACCATGGAGCC 180
QY 181 GGAAACAAGGACCATTTGCGTCACTTAAGGGTCTGTATTATCCAGATGTACCAATGTGAC 240
Db 181 GGAAACAAGGACCATTTGCGTCACTTAAGGGTCTGTATTATCCAGATGTACCAATGTGAC 240
QY 241 CAAGACCTCGTAGGCTGGCGCGCTCCCAAGGTGCCCTCAATTAACCACTGCACTTGC 300
Db 241 CAAGACCTCGTAGGCTGGCGCGCTCCCAAGGTGCCCTCAATTAACCACTGCACTTGC 300
QY 301 GGCTCTCGGACCTTTTACCTGTGTACAGGAGCAGCGCGGATGTCAATTCCTGTGTGCGCGA 360
Db 301 GGCTCTCGGACCTTTTACCTGTGTACAGGAGCAGCGCGGATGTCAATTCCTGTGTGCGCGA 360
QY 361 GGTGATGCGAGGGGAGCGCTGTTTGGCCCGCGGCTATCTCTTACTTGAAGGCTCTCG 420
Db 361 GGTGATGCGAGGGGAGCGCTGTTTGGCCCGCGGCTATCTCTTACTTGAAGGCTCTCG 420
```



PI Sallberg M;  
XX WPI: 2004-051480/05.  
DR P-PSDB; ADG47659.  
XX  
PT New purified or isolated nucleic acid useful for enhancing an immune  
response to a hepatitis C antigen comprises specific nucleotide sequences  
and the amino acid sequences.  
XX  
XX Example 1; SEQ ID NO 1; 83pp; English.  
XX  
CC The invention relates to a purified or isolated nucleic acid. The  
peptides are useful as immunogens for the treatment and prevention of  
hepatitis C virus (HCV) infection, in vaccine and immunogen compositions.  
CC The nucleic acid and the peptide enhance an immune response to a  
hepatitis C antigen and are potent immunogens. The present sequence is  
used in the exemplification of the invention.  
XX  
SQ Sequence 2061 BP; 427 A; 616 C; 571 G; 447 T; 0 U; 0 Other;  
  
Query Match 100.0%; Score 2061; DB 12; Length 2061;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 2061; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1 ATGGCGCCTATCACGGCCTATGCCAGCAGACAAAGGGGCTTTTGGGATGCATATCACC 60  
DB 1 ATGGCGCCTATCACGGCCTATGCCAGCAGACAAAGGGGCTTTTGGGATGCATATCACC 60  
  
QY 61 AGCTTGACCGCGCGGACAAAACCAAGGTGAGGGTGCAGATCGTGTCAACTGCT 120  
DB 61 AGCTTGACCGCGCGGACAAAACCAAGGTGAGGGTGCAGATCGTGTCAACTGCT 120  
  
QY 121 GCCCAGACTTCTTGCCACCTGCATTAAACGGGTGTGTGACTGTCTACCATGAGGC 180  
DB 121 GCCCAGACTTCTTGCCACCTGCATTAAACGGGTGTGTGACTGTCTACCATGAGGC 180  
  
QY 181 GGAAACAGGACCAATTCGTCACCTAAGGGTCTCTTTATCCAGATGTACCAATGTGGAC 240  
DB 181 GGAAACAGGACCAATTCGTCACCTAAGGGTCTCTTTATCCAGATGTACCAATGTGGAC 240  
  
QY 241 CAAGACCTGTAGGCTGGCCGCTCCCAAGGTGCCCGCTCATTAACCATGCACTTGC 300  
DB 241 CAAGACCTGTAGGCTGGCCGCTCCCAAGGTGCCCGCTCATTAACCATGCACTTGC 300  
  
QY 301 GGTCTCTCGACCTTTACCTGGTCAGAGCAGCGCGATGTCAATCTCTGCGCCGACGG 360  
DB 301 GGTCTCTCGACCTTTACCTGGTCAGAGCAGCGCGATGTCAATCTCTGCGCCGACGG 360  
  
QY 361 GGTGATGGCAGGGGAGCCTGTCTTCCGCCCGCCTATCTCTTACTTGAAGGCTCTCG 420  
DB 361 GGTGATGGCAGGGGAGCCTGTCTTCCGCCCGCCTATCTCTTACTTGAAGGCTCTCG 420  
  
QY 421 GGAGGCGCTCTGTGTGCGCGCAGACATGCGGTAGGCATATTACAGCCGCGGTATGC 480  
DB 421 GGAGGCGCTCTGTGTGCGCGCAGACATGCGGTAGGCATATTACAGCCGCGGTATGC 480  
  
QY 481 ACCCGTGGAGTGGCTAAGCGGTGGACTTATCCCGTAGAGCTTTAGAGCAACCATG 540  
DB 481 ACCCGTGGAGTGGCTAAGCGGTGGACTTATCCCGTAGAGCTTTAGAGCAACCATG 540  
  
QY 541 AGGTCCCGGTGTCTCAGACAACTCTCCCAACAGCAGTGTCCCGAGAGTACCAAGTG 600  
DB 541 AGGTCCCGGTGTCTCAGACAACTCTCCCAACAGCAGTGTCCCGAGAGTACCAAGTG 600  
  
QY 601 GCCCACTGATGCTCTCCACCGCAGCGGTAGAGCACCAGGTCCCGCGCATACGCA 660  
DB 601 GCCCACTGATGCTCTCCACCGCAGCGGTAGAGCACCAGGTCCCGCGCATACGCA 660  
  
QY 661 GCTCAGGGCTACAAAGTGTGTGTGCTCAACCCCTCGTGTGCAACATGGGCTTGGT 720  
DB 661 GCTCAGGGCTACAAAGTGTGTGTGCTCAACCCCTCGTGTGCAACATGGGCTTGGT 720  
  
QY 721 GCTTACATGTCCAAAGGCCCATGGGATTGTATCTTAAACATCAGGACTGGGGTGAGCAATT 780

DB 721 GCTTACATGTCCAAAGGCCCATGGGATTGTATCTTAAACATCAGGACTGGGGTGAGCAATT 780  
QY 781 ACTACTGGCAGCCGATCACGTTATTCACCTACCGCAAGTTCTTGGCGAGCGGGGTGT 840  
DB 781 ACTACTGGCAGCCGATCACGTTATTCACCTACCGCAAGTTCTTGGCGAGCGGGGTGT 840  
QY 841 TCAGGGGGTGTATGACATAATAATTTGTGACGAGTGCCTACCTCCACGATGCAACATCC 900  
DB 841 TCAGGGGGTGTATGACATAATAATTTGTGACGAGTGCCTACCTCCACGATGCAACATCC 900  
QY 901 ATCTTGGGCAATGGCACTGTCTTGAACCAAGCAGAGACCGCGGGGCGAGACTGTGTG 960  
DB 901 ATCTTGGGCAATGGCACTGTCTTGAACCAAGCAGAGACCGCGGGGCGAGACTGTGTG 960  
QY 961 CTGCGCACCGTACCCCTCCGGGCTCCGCTACTGTGCCCCATCTTAAACATCGAGGAGTT 1020  
DB 961 CTGCGCACCGTACCCCTCCGGGCTCCGCTACTGTGCCCCATCTTAAACATCGAGGAGTT 1020  
QY 1021 GCTCTGTCCACTACCGGAGAGATCCCTTTTATGGCAAGGCTATTCCCTTTGAAGCAATT 1080  
DB 1021 GCTCTGTCCACTACCGGAGAGATCCCTTTTATGGCAAGGCTATTCCCTTTGAAGCAATT 1080  
QY 1081 AAGGGGGGAGACATCTCATCTTGTGCCATCTCAAGAAAGATGCGAGCTGCCCGCA 1140  
DB 1081 AAGGGGGGAGACATCTCATCTTGTGCCATCTCAAGAAAGATGCGAGCTGCCCGCA 1140  
QY 1141 AAATCTGTGCGTGGGGTCAATGCGTGTCTACTACCGGGCTTGTATGTGTCCGTC 1200  
DB 1141 AAATCTGTGCGTGGGGTCAATGCGTGTCTACTACCGGGCTTGTATGTGTCCGTC 1200  
QY 1201 ATCCCGACAGTGTGACGTTGTCTGCTGGCACTGACGCGCTCATGACCGGCTTTACC 1260  
DB 1201 ATCCCGACAGTGTGACGTTGTCTGCTGGCACTGACGCGCTCATGACCGGCTTTACC 1260  
QY 1261 GCGCATCTCGATTGCGTGTATGACATGCAACAGTGTGTACCCAGACAGTGTGTGAGC 1320  
DB 1261 GCGCATCTCGATTGCGTGTATGACATGCAACAGTGTGTACCCAGACAGTGTGTGAGC 1320  
QY 1321 CTGACCTCTTACCTTACCAATGAGACATCAGCTTCCCGAGGATGTCTTCCGCTACT 1380  
DB 1321 CTGACCTCTTACCTTACCAATGAGACATCAGCTTCCCGAGGATGTCTTCCGCTACT 1380  
QY 1381 CAACGCTCGGGGTAGGACTGGCAGAGGAGCAGGCACTTACAGATTTGTGGCAACGGGG 1440  
DB 1381 CAACGCTCGGGGTAGGACTGGCAGAGGAGCAGGCACTTACAGATTTGTGGCAACGGGG 1440  
QY 1441 GAGCGTCTTCTGGCATGTTTGTGACTGTCTCTCTGCGAGTGTATGACGCGGGTGT 1500  
DB 1441 GAGCGTCTTCTGGCATGTTTGTGACTGTCTCTCTCTGCGAGTGTATGACGCGGGTGT 1500  
QY 1501 GCTTGTATGAGCTTACCGCCGAGACACAGTTAGGCTACGAGCATACATGAACACC 1560  
DB 1501 GCTTGTATGAGCTTACCGCCGAGACACAGTTAGGCTACGAGCATACATGAACACC 1560  
QY 1561 CCGGACTTCCGCTGTGCCAAGACCATCTTGAATTTTGGAGGGGCTCTTTACGGGTCTC 1620  
DB 1561 CCGGACTTCCGCTGTGCCAAGACCATCTTGAATTTTGGAGGGGCTCTTTACGGGTCTC 1620  
QY 1621 ACCCACTAGAGCCCATCTTCTATCCAGACAAAGCAGAGTGGGGAAAACTTCCCTAT 1680  
DB 1621 ACCCACTAGAGCCCATCTTCTATCCAGACAAAGCAGAGTGGGGAAAACTTCCCTAT 1680  
QY 1681 CTGTGTAGGCTACCAAGCCACCGTGTGCGTAGAGCTCAAGCCCTTCCCGCTGTGGAC 1740  
DB 1681 CTGTGTAGGCTACCAAGCCACCGTGTGCGTAGAGCTCAAGCCCTTCCCGCTGTGGAC 1740  
QY 1741 CAGATGTGGAAGTGTGATCCGCTCTCAAGCCACCCCTCCATGGGCCAACACTCTGTGTA 1800  
DB 1741 CAGATGTGGAAGTGTGATCCGCTCTCAAGCCACCCCTCCATGGGCCAACACTCTGTGTA 1800  
QY 1801 TATAGACTGGGGCTGTCTCCAGATGAAGTCAAGTGTGCGCAGCCAGTCCAGGATATATC 1860



Db 1801 TATAGACTGGCGCTGTCAGAAATGAAGTCAACCTGACGACCCAGTCAACCAAGTATATC 1860  
Qy 1861 ATGACATGATGTCGGCTGACCTGGAGGTGTCACGAGTACCTGGGTGCTGCTGGCGGC 1920  
Db 1861 ATGACATGATGTCGGCTGACCTGGAGGTGTCACGAGTACCTGGGTGCTGCTGGCGGC 1920  
Qy 1921 GTTCTGGCTGCTTTGGCCGGGATTCGCTATCCACAGGCTGGTGTCTAGTAGTAGG 1980  
Db 1921 GTTCTGGCTGCTTTGGCCGGGATTCGCTATCCACAGGCTGGTGTCTAGTAGTAGG 1980  
Qy 1981 ATTGTCTTGTCCGAAAGCCGGAATCATATACCCGAGGAAGTCTCTACCGGGAGTTC 2040  
Db 1981 ATTGTCTTGTCCGAAAGCCGGAATCATATACCCGAGGAAGTCTCTACCGGGAGTTC 2040  
Qy 2041 GATGAATGAAGAGTGTCTGA 2061  
Db 2041 GATGAATGAAGAGTGTCTGA 2061  
  
RESULT 5  
AAF83669  
ID AAF83669 standard; DNA; 6299 BP.  
XX AAF83669;  
AC AAF83669;  
XX  
DT 23-JUL-2001. (first entry)  
XX  
DE HCV NS34A ORF comprising pCMV-NS34A nucleic acid sequence.  
XX  
KW HCV; mutant; non-structural protein; NS; hepatitis C virus; mutation;  
KW catalytic domain; NS3; NS4; NS5; antiviral; vaccine; immunostimulant;  
KW immunotherapy; NS34A; ds.  
XX  
OS Synthetic.  
OS Hepatitis C virus.  
XX  
FH Key Location/Qualifiers  
FT CDS 1990..4050  
FT /\*tag= a  
FT /\*product= "NS34A"  
XX  
PN WO200138360-A2.  
XX  
XX 31-MAY-2001.  
PD  
XX 22-NOV-2000; 2000WO-US032326.  
PF  
XX 24-NOV-1999; 99US-0167502P.  
PR  
XX (CHIR ) CHIRON CORP.  
XX  
XX Coit D, Medina-Selby A, Selby M, Houghton M;  
XX WPI; 2001-343948/36.  
DR P-PSDB; AAB62633.  
XX  
PT Mutant non-structural (NS) Hepatitis C virus (HCV) polypeptide, useful as  
PT a vaccine against HCV, comprises a polypeptide having a mutation that  
PT functionally disrupts the catalytic domain of NS3.  
XX  
XX Disclosure; Fig 9; 340pp; English.  
XX  
XX The invention relates to an isolated mutant non-structural (NS) Hepatitis  
XX C virus (HCV) polypeptide, comprising a polypeptide having a mutation in  
XX the catalytic domain of NS3, where the mutation functionally disrupts the  
XX catalytic domain. The NS mutant polypeptides can include NS3, NS4 (NS4a  
XX and NS4b) NS5 (NS5a and NS5b) or portions thereof. The HCV polypeptide  
XX and polynucleotide (preferably DNA or a plasmid) compositions can be used  
XX in vaccines against HCV and as diagnostics. The antibodies raised against  
XX these polypeptides can also be used as diagnostics, or for passive  
XX immunotherapy. The antibodies are also useful for isolating and  
XX identifying HCV particles. The present sequence represents the nucleic  
XX acid sequence of the pCMV-NS34A comprising the ORF encoding a NS34A

CC polypeptide  
XX  
SQ Sequence 6299 BP; 1449 A; 1713 C; 1637 G; 1500 T; 0 U; 0 Other;  
Query Match 86.7%; Score 1787; DB 4; Length 6299;  
Best Local Similarity 91.7%; Pred No. 0;  
Matches 1889; Conservative 0; Mismatches 170; Indels 0; Gaps 0;  
Qy 1 ATGCGCGCTTATCACGGCTATGCCAGCAGACAAAGGGGCTTTTGGGATGCAATATCACC 60  
Db 1990 ATGGCGCCCATCACGGCTATGCCAGCAGACAAAGGGGCTTCTAGGGTGCAATATCACC 2049  
Qy 61 ACCTTGACCGGGGGGAGAAACACAGGTGGAGGTGAGGTTTCAGATCTGTCTCACTGCT 120  
Db 2050 AGCCTAACTGGCCGGGACAAAACCAAGTGGAGGGTGAGGTTCCAGATTGTGTCAACTGT 2109  
Qy 121 GCCCAGACTTTCTTGGCAACTGCAATTAACGGGGTGTGTGGACTGTCTACCATGAGAGC 180  
Db 2110 GCCCAAACTTCTTGGCAACTGCAATTAACGGGGTGTGTGGACTGTCTACCATGAGAGC 2169  
Qy 181 GGAACAAGACCAATTCGCTCACTTAAGGTCTCTTATCCAGATGTPACCAATGTGGAC 240  
Db 2170 GGAACGAGGACCATCGCGTCAACCAAGGGTCTCTCATCCAGATGTATACCAATGTAGAC 2229  
Qy 241 CAAGACCTGTTAGGCTGGCCCGCTCCCAAGGTGCCCGCTCATTAACACATGCATTCG 300  
Db 2230 CAAGACCTTGTGGGCTGGCCCGCTTCGCAAGGTACCCGCTCATTAACACATGCATTCG 2289  
Qy 301 GCCTCTCGGACCTTTACCTGTGTCAGGCGACCGCATGTCTATCTCTGCGCCGACCG 360  
Db 2290 GGCTCTCTCGGACCTTTACCTGTGTCAGGCGACCGCATGTCTATCTCTGCGCCGCGG 2349  
Qy 361 GGTGATGGCAGGGGAGCCCTGCTTTGCGCCCGCGCTATCTCTTACTTGAAGGCTCTCTG 420  
Db 2350 GGTGATAGCAGGGGAGCCCTGCTGTCGCCCGCGCCATTTCTACTTGAAGGCTCTCTG 2409  
Qy 421 GGAGGCTCTCTGTCGTCGCGCAGACATGCGCTGAGGATATTCAGAGCGCGGTATGC 480  
Db 2410 GGGGGTCCGCTGTGTGCGCCCGCGGCGACGCGCTGGGCATATTTAGGGCGCGGTGTG 2469  
Qy 481 ACCCGTGGAGTGGCTAAGCGGTGAGCTTTCATCCCGTAGAGAGCTTAGAGCAACCATG 540  
Db 2470 ACCCGTGGAGTGGCTAAGCGGTGAGCTTTCATCCCGTAGAGAGCTTAGAGCAACCATG 2529  
Qy 541 AGGTCCCCGGTGTTCCTCAGACAACTCTCCCAACAGCAGTCCCGCAGAGTACCAAGTG 600  
Db 2530 AGGTCCCCGGTGTTCCTCAGAGTAACTCTCTCCACAGTAGTGCCTCAGAGTTCAGGTG 2589  
Qy 601 GCCCAGCTGATGCTCCCAACCGGCGAGCGGTAAAGACCAAGGTCCCGCGCGCATACGCA 660  
Db 2590 GCTCACCTCCATGCTCCCAACAGGCGGCAAAAGCACCAAGGTCCCGCGCTGCATATGCA 2649  
Qy 661 GCTCAGGGCTACAGGTGCTGGTGTCTCAACCCCTCGCTGCTGCAACATGGGCTTGTGT 720  
Db 2650 GCTCAGGGCTATAGGTGCTAGTACTCAACCCCTCTGTTGCTGCAACATGGGCTTGTGT 2709  
Qy 721 GCTTACATGCTCAAGGCCCATCGGATTCATCAATCAGGACTGGGTGAGGACAAT 780  
Db 2710 GCTTACATGCTCAAGGGCTCATGGATCGATCTCAATCAGGACCGGGGTGAGACAAT 2769  
Qy 781 ACTACTGGCAGCCCGATCAGTATTCACCTACGGCAAGTTCCTTCCGACGCGGGGTGT 840  
Db 2770 ACCACTGGCAGCCCATCAGTACTCCACCTACGGCAAGTTCCTTCCGACGCGGGGTGT 2829  
Qy 841 TCAGGGGTGCTTATGACATATATTTGTGAGGTGCGCATCTCCACGATGCAACATCC 900  
Db 2830 TCAGGGGTGCTTATGACATATATTTGTGAGGTGCGCATCTCCACGATGCAACATCC 2889  
Qy 901 ATCTTGGGATTTGGCACTGTCTTGAACCAAGCAGACACCGCGGGGCGAGACTGACTGTG 960  
Db 2890 ATCTTGGGATTTGGCACTGTCTTGAACCAAGCAGACACTGCGGGGCGAGACTGCTGTG 2949  
Qy 961 CTCGCCACCGCTACCCCTCCGGGCTCCGTCACCTGTGCCCCCATCTAACATCGAGGAGTT 1020

|||||  
2950 CTCGCCACCGCCACCCCTCCGGCTCGTCACTGTGCCCTCCCAACATCGAGAGGTT 3009  
|||||  
1021 GCTCTGTCCACTACCGGAGAGATCCCTTTTATGGCAAGGCTATCCCTTGAAGCAATT 1080  
|||||  
3010 GCTCTGTCCACCACCGGAGAGATCCCTTTTATGGCAAGGCTATCCCTTGAAGCAATT 3069  
|||||  
1081 AAGGGGGGAGAGATCTCATCTTCTGCCACTCAAGAAAGAGTGGCAGAGCTCGCGCA 1140  
|||||  
3070 AAGGGGGGAGAGATCTCATCTTCTGCCACTCAAGAAAGAGTGGCAGAGCTCGCGCA 3129  
|||||  
1141 AAATCGTTCGCTGGCGTCAATCGGTGGCTTACTACCGCGCTTGAATGTCCTGTC 1200  
|||||  
3130 AAGCTGCTGCATTTGGGCATCAATGCGGTGGCTACTACCGCGCTTGAATGTCCTGTC 3189  
|||||  
1201 ATCCCGACAGTGGTGCAGCTGTGCTGCTGGCACTGACGCGCTCATGACCGGCTTACC 1260  
|||||  
3190 ATCCCGACAGCGGGGATGTTGCTGCTGGCACTGACGCGCTCATGACCGGCTTACC 3249  
|||||  
1261 GCGGACTTCGATTCGGGTGATAGACTCAACACAGTGTGTCAACCCAGACAGTTCAGC 1320  
|||||  
3250 GCGGACTTCGATTCGGGTGATAGACTCAACACAGTGTGTCAACCCAGACAGTTCAGC 3309  
|||||  
1321 CTTGACCTTACCTTACCAATGAGCAATCAGCTTCCCGAGGATGCTGTCTCCGTAAT 1380  
|||||  
3310 CTTGACCTTACCTTACCAATGAGCAATCAGCTTCCCGAGGATGCTGTCTCCGTAAT 3369  
|||||  
1381 CACGTCGGGTAGGACTGGCAGAGGAGCCAGGATCTACAGATTTGTGACCGGG 1440  
|||||  
3370 CACGTCGGGTAGGACTGGCAGAGGAGCCAGGATCTACAGATTTGTGACCGGG 3429  
|||||  
1441 GAGCGTCTTCTGGCATGTTGACTCGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1500  
|||||  
3430 GAGCGTCTTCTGGCATGTTGACTCGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3489  
|||||  
1501 GCTTGTATGAGCTTACGCCCGGAGACACAGTTAGGCTACGAGCATACATGACAC 1560  
|||||  
3490 GCTTGTATGAGCTTACGCCCGGAGACACAGTTAGGCTACGAGCATACATGACAC 3549  
|||||  
1561 CCGGACTTCCGTTGGTCAAGACCATCTTGAATTTGGAGGGGCTCTTTACGGGTCTC 1620  
|||||  
3550 CCGGACTTCCGTTGGTCAAGACCATCTTGAATTTGGAGGGGCTCTTTACGGGTCTC 3609  
|||||  
1621 ACCACATAGAGCCCACTTCTATCCAGACAGAGTGGGAAACCTTCCCTAT 1680  
|||||  
3610 ACTCATATAGTCCCACTTCTATCCAGACAGAGTGGGAAACCTTCCCTAT 3669  
|||||  
1681 CTGGTAGCGTACCAAGCCACCGTGTGCGCTAGAGCTCAAGCCCTTCCCGCTCGTGGAC 1740  
|||||  
3670 CTGGTAGCGTACCAAGCCACCGTGTGCGCTAGAGCTCAAGCCCTTCCCGCTCGTGGAC 3729  
|||||  
1741 CAGATGTGAAGTGTGATCGTCTCAAGCCCACTTCAAGCCCACTTCAAGCCCACTTCAAG 1800  
|||||  
3730 CAGATGTGAAGTGTGATCGTCTCAAGCCCACTTCAAGCCCACTTCAAGCCCACTTCAAG 3789  
|||||  
1801 TATAGACTGGGCTGTCCAGAAATGAAGTCAACCTGACGACCCAGTCAACCAAGTATATC 1860  
|||||  
3790 TACAGACTGGGCTGTCCAGAAATGAAGTCAACCTGACGACCCAGTCAACCAAGTATATC 3849  
|||||  
1861 ATGACATGATGTGCGCTGACCTGAGAGTGTGTCACGAGTACCTGGGTGTCTGTGGCGG 1920  
|||||  
3850 ATGACATGATGTGCGCTGACCTGAGAGTGTGTCACGAGTACCTGGGTGTCTGTGGCGG 3909  
|||||  
1921 GTTCTGGCTGTGTCGGGCTATGCTTATCCATACAGGCTGGTGTGATAGTATAGT 1980  
|||||  
3910 GTTCTGGCTGTGTCGGGCTATGCTTATCCATACAGGCTGGTGTGATAGTATAGT 3969  
|||||  
1981 ATTGTCTGTGCGGAAAGCGGCAATCATACCCGACAGGGAAGTCTCTACCGGAGTTC 2040  
|||||  
3970 GTTGTCTGTGCGGAAAGCGGCAATCATACCCGACAGGGAAGTCTCTACCGGAGTTC 4029  
|||||  
2041 GATGAATGAGAGTGTCT 2059  
|||||

Db 4030 GATGAGATGGAAGAGTGCT 4048  
RESULT 6  
AAN92106  
ID AAN92106 standard; DNA; 7310 BP.  
XX  
AC AAN92106;  
XX 09-SEP-2004 (revised)  
DT 25-MAR-2003 (revised)  
DT 02-MAR-1990 (first entry)  
XX  
DE Combined ORFs of the HCV cDNAs from clones K9-1 through 15e.  
XX Hepatitis C virus; HCV; non-A, non-B hepatitis; NANBH; ss.  
XX Hepatitis C virus.  
OS Unidentified.  
OS  
XX Key Location/Qualifiers  
FT CDS 3..7310  
/\*tag= a  
XX EP318216-A.  
XX 31-MAY-1989.  
XX 18-NOV-1989; 88EP-00310922.  
XX 18-NOV-1987; 87US-00122714.  
PR 30-DEC-1987; 87US-00139886.  
PR 26-FEB-1988; 88US-00161072.  
PR 06-MAY-1988; 88US-00191263.  
PR 26-OCT-1988; 88US-00263584.  
PR 14-NOV-1988; 88US-00271450.  
XX (CHIR ) CHIRON CORP.  
PA (CHIR ) CHIRON CORP.  
XX Houghton M, Choo QL, Kuo G;  
PI WPI; 1989-159274/22.  
DR P-PSDB; AAP2050.  
XX Purified hepatitis C virus - and associated nucleic acids and  
PT polypeptide(s).  
XX Claim 3; Fig 47-1 - 47-8; 139pp; English.  
XX It is a double-stranded nucleotide sequence of the open reading frame  
CC (ORF) (tag a) extending through clones K9-1 to 15e of hepatitis C virus  
CC (HCV) cDNA. It can be used to make oligomeric DNA hybridisation probes to  
CC detect the presence of HCV nucleic acids in samples. The polypeptide(s)  
CC it encodes could be used as immunoassay reagents and vaccines and to  
CC generate antibodies useful in diagnosis and passive immunotherapy for HCV  
CC infection/non-A, non-B hepatitis. (Updated on 25-MAR-2003 to correct PR  
CC field.) (Updated on 25-MAR-2003 to correct PI field.)  
CC Revised record issued on 09-SEP-2004 : Correction to keywords  
XX Sequence 7310 BP; 1491 A; 2217 C; 2058 G; 1540 T; 0 U; 4 Other;  
SQ  
Query Match 86.7%; Score 1786; DB 1; Length 7310;  
Best Local Similarity 91.7%; Pred. No. 0;  
Matches 1888; Conservative 0; Mismatches 170; Indels 0; Gaps 0;  
QY 2 TGGCGCCCTATCACGGCCTATGCCAGCAGACAAAGGGGCTTTTGGGATGCATAATCACCA 61  
Db 1729 TGGCGCCCTATCACGGCCTATGCCAGCAGACAAAGGGGCTTTTGGGATGCATAATCACCA 1788  
QY 62 GCTTGACCGCGCGGACAAAACCAAGGTGGAGGTTCAGATCGTGTCACTGCTG 121  
|||||

Db 1789 GCCTAAGCTGGCGGACAAAAACAAGTGGAGGGTGAGGTCAGATTGTGTCAACTGCTG 1848  
Qy 122 CCCAGACTTTCTGGCAACCTGCATTAAACGGGGGTGTGTGAGACTGTCTACCATGAGCCG 181  
Db 1849 CCCAAACCTTCTGGCAACGTGCATCAATGGGGTGTGCTGACTGTCTACCAAGGGCGG 1908  
Qy 182 GAACAGGACCATTTGCGTCACTTAAGGGTCTGTATTCAGATGTACCAATGTGGACC 241  
Db 1909 GAACGAGGACCATCGCGTCAACCAAGGGTCTGTATTCAGATGTATACCAATGTAGACC 1968  
Qy 242 AAGACCTCGTAGGTGGCGCGCTCCCAAGGTGCGCTCATTAACACCATGACATTTGG 301  
Db 1969 AAGACCTTGTGGGTGGCGCGCTCCGCAAGGTAGCGCTCATTAACACCATGACATTTGG 2028  
Qy 302 GCTCTCGGACCTTTACCTGTCACGAGGACGCGGATGTCAATTCCTGTGCGCGGACGG 361  
Db 2029 GCTCTCGGACCTTTACCTGTCACGAGGACGCGGATGTCAATTCCTGTGCGCGGCGG 2088  
Qy 362 GTGATGGCAGGGGAGCCTGCTTTGCGCGCGGCTATCTCTTAATTGAAGGTCTCTCGG 421  
Db 2089 GTGATAGCAGGGGAGCCTGCTGTGCGCGCGGCTATTTAGGGCGCGGCTCTCGG 2148  
Qy 422 GAGGCTCTGCTGTGCGCGGACGAGCATGCGGTAGGCATATTCAGAGCGCGGATGCA 481  
Db 2149 GGGGTCCGCTGTTGTGCGCGCGGCGACGCGGTGGGCATATTTAGGGCGCGGCTGTGCA 2208  
Qy 482 CCGGTGAGTGGGTAAAGCGGTGGACTTCATCCCGTAGAGAGCTTAGAGACAACCATGA 541  
Db 2209 CCGGTGAGTGGGTAAAGCGGTGGACTTCATCCCGTAGAGAGCTTAGAGACAACCATGA 2268  
Qy 542 GGTCTCCCGGTGTTCTCAGACAACCTCTCCCAAGCAGAGTGCCTCCAGAGTACCAAGTGG 601  
Db 2269 GGTCTCCCGGTGTTCAAGGATAAATCTCTCTCCACAGTAGTGCCCGAGAGCTTCCAGGTGG 2328  
Qy 602 CCACCTGATGCTCCACCGCGAGCGGTAAAGAGCACAGAGTCCCGGCGGATACGCAG 661  
Db 2329 CTCACCTCCATGCTCCACAGGAGGAGCGGCAAAAGCACAAAGTCCCGGCTGCAATGCG 2388  
Qy 662 CTCAGGCTCAAGGTGCTGCTCAACCCCTCCGTTGTGCAACAATGGGCTTTGGTG 721  
Db 2389 CTCAGGCTCAAGGTGCTGCTCAACCCCTCTGTTGCTGCAACAATGGGCTTTGGTG 2448  
Qy 722 CTTCATGTCCTAAGGCTCATGGATTCATCAATCAGGACTGGGGTGAGGACAATTA 781  
Db 2449 CTTCATGTCCTAAGGCTCATGGATTCATCAATCAGGACTGGGGTGAGGACAATTA 2508  
Qy 782 CTACTGCGAGCCGATCAGCTATTCACCTACGGCAAGTTCCTTGGCGGCGGGTGT 841  
Db 2509 CCACTGCGAGCCGATCAGCTATTCACCTACGGCAAGTTCCTTGGCGGCGGGTGT 2568  
Qy 842 CAGGGGTGCTTATGACATAATTTGTGACAGTSCCACTCCACGGATGCAACATCCA 901  
Db 2569 CAGGGGTGCTTATGACATAATTTGTGACAGTSCCACTCCACGGATGCAACATCCA 2628  
Qy 902 TCTTGGCATTTGGCACTGTCTTTGACCAAGCAGAGACCGGGGGGCGAGACTGACTGTGC 961  
Db 2629 TCTTGGCATTTGGCACTGTCTTTGACCAAGCAGAGACTGCGGGGGCGAGACTGTTGTGC 2688  
Qy 962 TCGCCACCGCTACCCCTCGGGTCTCGTCACTGTGCCCCATCTTAACATCGAGAGGTG 1021  
Db 2689 TCGCCACCGCCACCCCTCGGGTCTCGTCACTGTGCCCCATCTTAACATCGAGAGGTG 2748  
Qy 1022 CTCGTGCTCACTACGGAGATCCCTTTTATGGCAGGCTATTCCTCTTGAAGCAATTA 1081  
Db 2749 CTCGTGCTCACTACGGAGATCCCTTTTATGGCAGGCTATTCCTCTTGAAGCAATTA 2808  
Qy 1082 AGGGGGGAGACATCTCATCTTCTGCGCACTCAAGAGAGAGTGGCAGGCTCGCGCAA 1141  
Db 2809 AGGGGGGAGACATCTCATCTTCTGCTCATTAAGAGAGAGTGGCAGGCTCGCGCAA 2868  
Qy 1142 AACTGTGCGGTGGGGTCAATGCGGCTTACTACCGCGGCTTGTGTGCTCGGTCA 1201  
Db 2869 AGCTGTGCGATTGGGCATCAATGCGGCTTACTACCGGCTTGTGAGTGTGCTCGGTCA 2928

Qy 1202 TCCCGACCAAGTGGTACGTTGTGTCGTGGCAACTGACGCCCTCATGACCGGCTTTACCG 1261  
Db 2929 TCCCGACCAAGCGCGATGTTGTGTCGTGGCAACGATGCCCTCATGACCGGCTATACCG 2988  
Qy 1262 GCGACTTCATTCCGTGTAGTAGTGCACACAGTGTGTACCCAGACAGTGCAGTTCAGCC 1321  
Db 2989 GCGACTTCGACTCGGTGTAGTAGTGCACATACGTTGTGTCTACCCAGACAGTGCAGTTCAGCC 3048  
Qy 1322 TTGACCTTACCTTACCATTTGAGACAATCAGCTTCCCGAGGATGCTGTCTCCGCTACTC 1381  
Db 3049 TTGACCTTACCTTACCATTTGAGACAATCAGCTTCCCGAGGATGCTGTCTCCGCTACTC 3108  
Qy 1382 AACGTGCGGGTAGGACTGGCAGAGGAAGCCAGGCACTACAGATTGTGGCACCGGGGG 1441  
Db 3109 AACGTGCGGGTAGGACTGGCAGAGGAAGCCAGGCACTACAGATTGTGGCACCGGGGG 3168  
Qy 1442 AGGTCCTTCTGCGATGTTTGTAGTCTGTCGTCTCTGCGAGTGTCTATGACGCGGGTGTG 1501  
Db 3169 AGCGCCCTTCCGCGATGTTTGTAGTCTGTCGTCTCTGTCGTGTGTATGACGAGGCTGTG 3228  
Qy 1502 CTTGTGTATGAGTTCACGCCCGCGGAGACTACAGTTAGGCTACGAGGCTACATGAACACCC 1561  
Db 3229 CTTGTGTATGAGTTCACGCCCGCGGAGACTACAGTTAGGCTACGAGGCTACATGAACACCC 3288  
Qy 1562 CGGGACTTCCCGGTGTCGCAAGACCATCTTGAATTTTGGGAGGGCGTCTTTACGGGTCTCA 1621  
Db 3289 CGGGGTCTCCCGGTGTCGCAAGACCATCTTGAATTTTGGGAGGGCGTCTTTACAGGCTCA 3348  
Qy 1622 CCACATAGACGCCCACTTCTATCCAGACAAGCAGAGTGGGGGAAAACCTTCCCTATC 1681  
Db 3349 CTCATATAGATGCCCACTTCTATCCAGACAAGCAGAGTGGGGGAAAACCTTCCCTATC 3408  
Qy 1682 TGTAGCTGACCAAGCCCGGTGCGCTAGAGCTCAAGCCCTCCCGCTCGTGGGAGC 1741  
Db 3409 TGTAGCTGACCAAGCCCGGTGCGCTAGAGCTCAAGCCCTCCCGCTCGTGGGAGC 3468  
Qy 1742 AGATGTGGAAGTGTCTGATCCGCTCTCAAGCCCAACCTCCATGGGCGCAACACCTCTGCTAT 1801  
Db 3469 AGATGTGGAAGTGTCTGATCCGCTCTCAAGCCCAACCTCCATGGGCGCAACACCTCTGCTAT 3528  
Qy 1802 ATAGACTGGGCGTGTCCAGAAATGAAGTCAACCTGACCAAGTGCAGTGCAGGATATCA 1861  
Db 3529 ACAGACTGGGCGTGTTCAGAAATGAATCAACCTGACGCAACCCAGTCAACAAATACATCA 3588  
Qy 1862 TGACATGATGTGCGTGTGAGTGTGTCAGAGTGTCAAGTGTGCTCGTGGGCGG 1921  
Db 3589 TGACATGATGTGCGGCGGACCTGGAGTGTCAAGGACCTGGGTGCTCGTGGGCGG 3648  
Qy 1922 TTCTGGCTGCTTTGGCGGCTATTTGCTATCCACAGGCTGCGTGTGCTCATAGTAGGAGGA 1981  
Db 3649 TTCTGGCTGCTTTGGCGGCTATTTGCTATCCACAGGCTGCGTGTGCTCATAGTAGGAGGA 3708  
Qy 1982 TTGCTGTGTCGGAAGCGGCAATCATACCGGACAGGAAAGTCTCTTACCGGAGGTTCG 2041  
Db 3709 TGCTGTGTCGGAAGCGGCAATCATACCTGACAGGAAAGTCTCTTACCGGAGGTTCG 3768  
Qy 2042 ATGAATGGAAGAGTCT 2059  
Db 3769 ATGAGATGGAAGAGTCT 3786

## RESULT 7

AAN90336

ID AAN90336 standard; DNA; 7310 BP.

XX

AAN90336;

XX

25-MAR-2003 (revised)

19-JUL-2001 (revised)

01-NOV-1989 (first entry)

XX

DE Composite hepatitis C virus (HCV) cDNA.

XX Hepatitis C virus; cDNA; clone 15e; clone k9-1; probe; vaccine; ds.  
 KW Pan troglodytes.  
 XX GB2212511-A.  
 XX 26-JUL-1989.  
 XX 18-NOV-1988; 88GB-00027024.  
 XX 18-NOV-1987; 87US-00122714.  
 PR 30-DEC-1987; 87US-00139886.  
 PR 26-FEB-1988; 88US-00161072.  
 PR 26-OCT-1988; 88US-00263584.  
 XX (CHIR ) CHIRON CORP.  
 PA Houghton M, Choo QL, Kuo G;  
 XX WPI; 1989-215054/30.  
 DR P-PSDB; AAP90288.  
 XX Hepatitis C virus gene - used for prodn. of polynucleotide probes  
 PT polypeptide(s) and antibodies for diagnosis, prevention and treatment of  
 PT infection.  
 XX Disclosure; Fig 47; 30pp; English.  
 XX The sequence shows a composite hepatitis C virus (HCV) cDNA, derived by  
 CC aligning clones k9-1 through 15e in 5'-3' direction. The cDNA encodes  
 CC antigens which react with antibodies in patients with non-A non-B  
 CC hepatitis (NANBH). The cDNA can be used to design probes, or to  
 CC synthesise polypeptides, which are used to diagnose HCV-induced NANBH, to  
 CC raise antibodies for immunoassay or treatment, or to produce vaccines.  
 CC See also AAP90288, and AAN90303-35. (N.B. This record was resubmitted to  
 CC correct errors in the sequence.) (Updated on 25-MAR-2003 to correct PR  
 CC field.)  
 XX Sequence 7310 BP; 1495 A; 2218 C; 2058 G; 1539 T; 0 U; 0 Other;  
 SQ

Query Match 86.7%; Score 1786; DB 1; Length 7310;  
 Best Local Similarity 91.7%; Pred. No. 0;  
 Matches 1888; Conservative 0; Mismatches 170; Indels 0; Gaps 0;

QY 2 TGGCGCTATACGGGCTATGCCAGCAGACAAAGGGGCTTTGGATGCATAATCACCA 61  
 DB 1729 TGGCGCCCATCACGGGTACGCCAGACAAAGGGGCTCTCTAGGGTGCATAATCACCA 1788

QY 62 GCTTGACCGCGGGACAAAACAGGTGGAGGTGAGGTTCAGATCGTGTCACCTGCTG 121  
 DB 1789 GCTTAAGTGGCGGGGACAAAACAGGTGGAGGTGAGGTTCAGATCGTGTCACCTGCTG 1848

QY 122 CCAGACTTTCTTTGGCAACTGCATTAACGGGGTGTGTTGGACTGTCTACCATGGAGCCG 181  
 DB 1849 CCCAACTTCTTGGCAACTGCATTAACGGGGTGTGTTGGACTGTCTACCATGGAGCCG 1908

QY 182 GAACAGGACCATTTGGTCACTAAGGTCTCTGTTATCCAGATGTACCAATGTGGACC 241  
 DB 1909 GAACAGGACCATTTGGTCACTAAGGTCTCTGTTATCCAGATGTACCAATGTGGACC 1968

QY 242 AAGACTCTGTAGGCTGGCGGCTCCCAAGGTGCCCGCTCATTAACACCATGCATTTGG 301  
 DB 1969 AAGACTCTGTAGGCTGGCGGCTCCCAAGGTGCCCGCTCATTAACACCATGCATTTGG 2028

QY 302 GCTCTCTCGGACCTTTACCTGGTCAAGGACACCGCATGTCAFTCTGTGCGCCGACGG 361  
 DB 2029 GCTCTCTCGGACCTTTACCTGGTCAAGGACACCGCATGTCAFTCTGTGCGCCGACGG 2088

QY 362 GTGATGGAGGGGACGCTGCTTTCCCGCGGCTATCTTACTTTGAAAGGCTCTCTCG 421  
 DB 2089 GTGATGGAGGGGACGCTGCTTTCCCGCGGCTATCTTACTTTGAAAGGCTCTCTCG 2148

QY 422 GAGGCCCTCTGTGTGCCCGCAGGACATGCCGTAGGCATATTCAGAGCCGCGGTATGCA 481  
 DB 2149 GGGGTCCGTGTGTGTGCCCGCGGGGACGCGGTGGGCATATTTAGGGCCGCGGTGCA 2208

QY 482 CCCGTGGAGTGGCTAAGGCGGTGGACTTTCATCCCCTAGAGAGCTTAGAGACAACATGA 541  
 DB 2209 CCCGTGGAGTGGCTAAGGCGGTGGACTTTCATCCCCTAGAGAGCTTAGAGACAACATGA 2268

QY 542 GGTCCCGGTGTCTCAGACAACCTCTCCCAACAGCAGTGTCCCGAGAGCTACCAAGTGG 601  
 DB 2269 GGTCCCGGTGTCTCAGACAACCTCTCCCAACAGCAGTGTCCCGAGAGCTACCAAGTGG 2328

QY 602 CCCACTGCATCTCTCCCAACAGCAGTGTCCCGAGAGCTTAGAGACAACATGA 661  
 DB 2329 CTCACCTCCATGCTCTCCCAACAGCAGTGTCCCGAGAGCTTAGAGACAACATGA 2388

QY 662 CTCAGGGCTACAAGGTGTGTGTCTCAACCCCTCCGTTCCTGCTCAACATTAAGTGGTGTG 721  
 DB 2389 CTCAGGGCTACAAGGTGTGTGTCTCAACCCCTCCGTTCCTGCTCAACATTAAGTGGTGTG 2448

QY 722 CTTCATGTCCAAAGGCCCATGGATGTATCTTAAACATCAGGACTGGGTGAGGACAATTA 781  
 DB 2449 CTTCATGTCCAAAGGCCCATGGATGTATCTTAAACATCAGGACTGGGTGAGGACAATTA 2508

QY 782 CTACTGGCAGCCCGCATCAGTATTCACCTACGCAAGTTCCTTGCAGACGCGGGTGT 841  
 DB 2509 CCACCTGGCAGCCCGCATCAGTATTCACCTACGCAAGTTCCTTGCAGACGCGGGTGT 2568

QY 842 CAGGGGTGTCTTATCACAATAATTTGTGACGATGTCCATCTCCACGGATGCAACATCCA 901  
 DB 2569 CAGGGGTGTCTTATCACAATAATTTGTGACGATGTCCATCTCCACGGATGCAACATCCA 2628

QY 902 TCTTGGGCACTGGCACTGCTTGTACCAAGCAGAGACCGCGGGGGGAGACTGACTGTGC 961  
 DB 2629 TCTTGGGCACTGGCACTGCTTGTACCAAGCAGAGACCGCGGGGGGAGACTGACTGTGC 2688

QY 962 TCGCCACCGCTACCCCTCCGGCTCCGTCACTGTGCCCATCTTAACATCGAGGAGTGTG 1021  
 DB 2689 TCGCCACCGCTACCCCTCCGGCTCCGTCACTGTGCCCATCTTAACATCGAGGAGTGTG 2748

QY 1022 CTCTGTCTCACTACCGAGAGATCCCTTTTATGGCAAGCTATTCCTTGAAGCAATTA 1081  
 DB 2749 CTCTGTCTCACTACCGAGAGATCCCTTTTATGGCAAGCTATTCCTTGAAGCAATTA 2808

QY 1082 AGGGGGGAGACATCTCTCTCCCACTCAAGAGAGAGTCCGACGACTCGCGCAA 1141  
 DB 2809 AGGGGGGAGACATCTCTCTCCCACTCAAGAGAGAGTCCGACGACTCGCGCAA 2868

QY 1142 AACTGGTGGCTGGCGCTCAATGCGGTCTTACTACCGCGCTTGTGTGTCCGTCA 1201  
 DB 2869 AGCTGGTGGCATTTGGGCATCATGCGGTGCTTACTACCGCGCTTGTGTGTCCGTCA 2928

QY 1202 TCCCGACAGTGGTGAAGTGTGTGTGGGCAACTGAGCCCTCATGAGCCGGTGTACCG 1261  
 DB 2929 TCCCGACAGTGGTGAAGTGTGTGTGGGCAACTGAGCCCTCATGAGCCGGTGTACCG 2988

QY 1262 GCGACTTCGATTCGTTGATAGTCAACACCTGTGTCAACAGAGTGCAGCTTCAGCC 1321  
 DB 2989 GCGACTTCGACTTCGTTGATAGTCAACACCTGTGTCAACAGAGTGCAGCTTCAGCC 3048

QY 1322 TTGACCTTACCTTACCATTTAGACAATCACCGCTTCCCGAGGATGTGTCTCCGCTAC 1381  
 DB 3049 TTGACCTTACCTTACCATTTAGACAATCACCGCTTCCCGAGGATGTGTCTCCGCTAC 3108

QY 1382 AACGTGGGGTGAAGCTGGCAGAGAGGAGCAGGATCTACAGATTTGTGGACCGGGGG 1441  
 DB 3109 AACGTGGGGTGAAGCTGGCAGAGGAGCAGGATCTACAGATTTGTGGACCGGGGG 3168

QY 1442 AGCGTCTCTTCGCAATGTTTGTGCTCTCTGCGAGTGTATGACGCGGGTGTG 1501  
 DB 3169 AGCGTCTCTTCGCAATGTTTGTGCTCTCTGCGAGTGTATGACGAGGCTGTG 3228

QY 1502 CTTGGTATGAGCTTACGGCCCGCGAGACCAAGCTTAGGCTACGAGCATATCAACACCC 1561

Db 3229 CTTGGTATGAGCTCAGCCCGCGAGACTACGTTAGGCTACAGCGTACATGAACACCC 3288  
QY 1562 CGGAGCTTCCGCTGTCAGAGACATCTTGAATTTTGGAGGCGCTTTTACGGGCTCA 1621  
Db 3289 CGGGCTTCCGCTGTCAGAGACATCTTGAATTTTGGAGGCGCTTTTACAGGCTCA 3348  
QY 1622 CCACATAGAGCGCCACTTCTTATCCAGACAAAGCAGAGTGGGAAACCTTCCCTATC 1681  
Db 3349 CTCATATAGATGCCACTTTCTATCCAGACAAAGCAGAGTGGGAAACCTTCCCTACC 3408  
QY 1682 TGTAGGCTACCAAGCCACGCTGTGCTAGAGCTCAAGCCCTCCCGCTGTGGAGCC 1741  
Db 3409 TGTAGGCTACCAAGCCACGCTGTGCTAGGCTCAAGCCCTCCCGCTGTGGAGCC 3468  
QY 1742 AGATGTGGAAGTCTTGATCCGCTCAAGCCACCTTCCATGGGCCCAACACCTCTCTAT 1801  
Db 3469 AGATGTGGAAGTCTTGATCCGCTCAAGCCACCTTCCATGGGCCCAACACCTCTCTAT 3528  
QY 1802 ATAGACTGGCGCTGTCCAGAAATGAAGTCAACCTGACGACCCAGTCAACAAATATCA 1861  
Db 3529 ACAGACTGGCGCTGTTTCAAGTGAATCAACCTGACGACCCAGTCAACAAATATCA 3588  
QY 1862 TGACATGTATGCTGGCTGACCTGAGTGTCTACGAGTACCTGGGTGCTGTGGCGGG 1921  
Db 3589 TGACATGCTATGCTGGCGGACCTGGAGTGTCTACGAGCACTGGGTGCTGTGGCGGG 3648  
QY 1922 TTCTGGCTGTTTGGCGGCTATGCTTCCACAGCTGCGGTGCTAGTGTAGGA 1981  
Db 3649 TCTGGCTGTTTGGCGGCTATGCTTCCACAGCTGCGGTGCTAGTGTAGGA 3708  
QY 1982 TTCTGTTGTCGGAAAGCCGCAATCATACCCGACAGGGAAGTCTCTACCGGGAGTTCG 2041  
Db 3709 TCGTCTGTTCCGGAAGCCGCAATCATACCTGACAGGGAAGTCTCTACCGAGAGTTCG 3768  
QY 2042 ATGAATGGAAGTGTCT 2059  
Db 3769 ATGAGATGGAAGTGTCT 3786

RESULT 8  
AAQ98221 standard; cDNA to mRNA; 7310 BP.

XX  
AC  
XX  
DT 25-MAR-2003 (revised)  
DT 15-AUG-1996 (first entry)  
XX  
DE Hepatitis C virus clone genome.  
XX  
KW Hepatitis C virus; HCV; antigen; detection; diagnosis; vaccine;  
KW antibodies; immunoprophylaxis; sera; serum; ds.  
XX  
OS Hepatitis C virus.  
XX  
PN US5443965-A.  
XX  
PD 22-AUG-1995.  
XX  
PF 05-APR-1991; 91US-00681703.  
XX  
PR 06-APR-1990; 90US-00505611.  
PR 09-OCT-1990; 90US-00594854.  
XX  
PA (GENE-) GENELABS INC.  
XX  
PI Moeckli R, Reyes GR, Kim JP;  
XX  
XX WPI; 1995-302120/39.  
XX

PT New nucleic acids encoding hepatitis C virus antigens - used to develop  
PT prods. for detection of HCV-infected sera and prodn. of vaccines and anti

PT -HCV antibodies.  
XX Example 4; Fig 11; 71pp; English.  
XX Hepatitis C virus (HCV) antigens can be used for detecting HCV infected  
CC sera and individuals infected with HCV. They can also be used in an anti-  
CC HCV vaccine or for the production of anti-HCV antibodies which can be  
CC used for passive immunoprophylaxis. The antigens consistently identify  
CC more HCV positive serum samples with a high degree of specificity. See  
CC AAQ98202-14 and AAR81939-51. (Updated on 25-MAR-2003 to correct Pf  
CC field.) (Updated on 25-MAR-2003 to correct PR field.)  
XX

SQ Sequence 7310 BP; 1494 A; 2217 C; 2060 G; 1539 T; 0 U; 0 Other;

Query Match 86.7%; Score 1786; DB 2; Length 7310;  
Best Local Similarity 91.7%; Pred No. 0;  
Matches 1888; Conservative 0; Mismatches 170; Indels 0; Gaps 0;  
QY 2 TGGCGCCTATCACGCGCTATGCCAGCAGACAAAGGGCGCTTTGGGATGCATAATCACCA 61  
Db 1729 TGGCGCCTATCACGCGCTATGCCAGCAGACAAAGGGCGCTTTGGGATGCATAATCACCA 1788  
QY 62 GCTTACACCGCGCGGACAAAACACAGGTGGAGGTGAGTTTCAGATCGTGTCACTGCTG 121  
Db 1789 GCTTAACTGGCGCGGACAAAACCAAGTGGAGGTGAGTCCAGATTGTCACTGCTG 1848  
QY 122 CCCAGACTTTCTTGGCAACCTGTCATTAACGGGGTGTGTGGACTGTCTACCATGGAGCG 181  
Db 1849 CCCAAACCTTCTTGGCAACCTGTCATTAAGTGGGTGTGTGGACTGTCTACCAAGGGCG 1908  
QY 182 GAACAAAGGACCAATTCGCTCACCTAAGGGTCTGTGTATCCAGATGTACACCAATGTGAC 241  
Db 1909 GAACGAGGACCAATTCGCTCACCAAGGGTCTGTGTATCCAGATGTATACCAATGTAGAC 1968  
QY 242 AAGACCTGTAGCTGTGGCGCGCTCCCAAGGTGCCCGCTCATTAACACCATGCATGTCG 301  
Db 1969 AAGACCTGTGTGGCGCTGTGGCGCGCTCCCAAGGTAGCGCTCATTTGACACCCCTGAC 2028  
QY 302 GCTCTCTCGGACCTTTTACCTGTGTACGAGGACGCGGATGTCAATCTGTGCGCGGACGG 361  
Db 2029 GCTCTCTCGGACCTTTTACCTGTGTACGAGGACGCGGATGTCAATCTGTGCGCGGCGG 2088  
QY 362 GTGATGGCAGGGCGACCTGTCTTTCGCGCGCGCTATCTCTTACTTTGAAAGCTCTCTCG 421  
Db 2089 GTGATAGCAGGGCGACCTGTGTGCGCGCGCGCTATCTCTTACTTTGAAAGCTCTCTCG 2148  
QY 422 GAGGCGCTGTGTGTGCGCGCGGACGACATGCGGTAGGCAATATTCAGAGCGCGGTATGCA 481  
Db 2149 GGGGTCCGCTGTGTGCGCGCGGACGCGGTGGGCAATATTTAGGCGCGCGGTGTGCA 2208  
QY 482 CCGGTGGAGTGGCTAAGGGGTGGGACTTTCATCCCGTAGAGCTTAGAGCAACCATGA 541  
Db 2209 CCGGTGGAGTGGCTAAGGGGTGGGACTTTCATCCCGTAGAGCTTAGAGCAACCATGA 2268  
QY 542 GGTCCCGCGGTGTCTCAGACAACTCCTCCCAAGCAGAGTGCCTCCAGAGCTTACCAAGTGG 601  
Db 2269 GGTCCCGCGGTGTCTCAGAGTAACTCTCTCCAGAGTGTGCGCGCGGCTTCCAGGTGG 2328  
QY 602 CCCACCTGCATGCTCCCAACCGGACGCGGTAAAGACCAAGGTCCCGCGCGGTATGCGAG 661  
Db 2329 CTCACCTCCATGCTCCCAACGAGCGGCAAAAGCACAAGGTCCCGGCTGCATATGCA 2388  
QY 662 CTCAGGGCTACAGGTGTGTGTCTCAACCGCTCCGTGTGCTGCAACATGGCTTTGGTG 721  
Db 2389 CTCAGGGCTATAGGTGTGTGTCTCAACCGCTCCGTGTGCTGCAACATGGCTTTGGTG 2448  
QY 722 CTTACATGTCCAAAGGCCCATGGATTTGATCTTAACATCAGGACTGGGTGAGGACAAATTA 781  
Db 2449 CTTACATGTCCAAAGGCTCATGGATTCGATCTTAACATCAGGACCGGGGTGAGAACAA 2508  
QY 782 CTACTGGAGCGCCGATCATCGTATTCACCTACGGCAAGTTCCTTCCGACGCGGGTGT 841  
Db 2509 CCACCTGGCAGCCCATCATCGTACTCCACCTACGGCAAGTTCCTTCCGACGCGGGTGT 2568

QY 842 CAGGGGCTTATGACATAAATTTGTGACAGTGCACCTCCACGGATGCAATCA 901  
DB 2569 CGGGGGCGCTTATGACATAAATTTGTGACAGTGCACCTCCACGGATGCAATCA 2628  
QY 902 TCTTGGGCAATGGCACTGTCCTTTGACAAAGCAGAGACCGCGGGGCGAGACTGCTGTC 961  
DB 2629 TCTTGGGCAATGGCACTGTCCTTTGACAAAGCAGAGACTCGGGGGCGAGACTGCTGTC 2688  
QY 962 TCGCCACCGTACCCCTCCGGGCTCCGTCACGTGTCCTTAACTAATCGAGAGGTTG 1021  
DB 2689 TCGCCACCGTACCCCTCCGGGCTCCGTCACGTGTCCTTAACTAATCGAGAGGTTG 2748  
QY 1022 CTCTGTCCACTACCGGAGATCCCTTTTATGGCAAGCTATTCCTTTGAAGCAATTA 1081  
DB 2749 CTCTGTCCACTACCGGAGATCCCTTTTATGGCAAGCTATTCCTTTGAAGCAATTA 2808  
QY 1082 AGGGGGGAGACATCATCTTCTGCACTCAAGAAAGTGCAGACGCTCGCGCAA 1141  
DB 2809 AGGGGGGAGACATCATCTTCTGCACTCAAGAAAGTGCAGACGCTCGCGCAA 2868  
QY 1142 AACTGTGCGTGTGGCGTCAATATCCGTCGCTTATACCGGGCTTGTGATGTGCTCA 1201  
DB 2869 AGTGTGCGATTTGGGCATCAATGCGCTGCTTACTACCGCGCTTGTGACGTGCTGCA 2928  
QY 1202 TCCGACACGATGCTGACGTTGCTGTCGTCGCACTGACCGCTCATGACCGGCTTTACCG 1261  
DB 2929 TCCGACACGATGCTGACGTTGCTGTCGTCGCACTGACCGCTCATGACCGGCTTTACCG 2988  
QY 1262 GCGACTTCGATTCGCTGATAGACTGCAACCGTGTGTACCCAGACAGCTCGACTTCAGCC 1321  
DB 2989 GCGACTTCGATTCGCTGATAGACTGCAACCGTGTGTACCCAGACAGCTCGACTTCAGCC 3048  
QY 1322 TTGACCTCTACCTTCACTATGACAAATCAOGCTTCCCGAGTGTGTCTCCGTACTC 1381  
DB 3049 TTGACCTCTACCTTCACTATGACAAATCAOGCTTCCCGAGTGTGTCTCCGTACTC 3108  
QY 1382 AAGTCGGGGTAGGACTGCGAGGGAAGCCAGGCACTACAGATTGTGGCAGCGGGG 1441  
DB 3109 AAGTCGGGGTAGGACTGCGAGGGAAGCCAGGCACTACAGATTGTGGCAGCGGGG 3168  
QY 1442 AGCGTCTCTTGTGCACTGTTGACTCGCTCTCTCTGCGAGTGTGTATGACGGGGTTGTG 1501  
DB 3169 AGCGTCTCTTGTGCACTGTTGACTCGCTCTCTCTGCGAGTGTGTATGACGGGGTTGTG 3228  
QY 1502 CTTGTATGACTTACCGCCGCGAGACCAAGTGTAGGCTACGAGCATACATGAACACC 1561  
DB 3229 CTTGTATGACTTACCGCCGCGAGACTACAGTTAGGCTACGAGCGTACATGAACACC 3288  
QY 1562 CGGACTTCCGCTGTCGAAGACCATCTTGAATTTTGGGAGGCGTCTTTACGGGTCTCA 1621  
DB 3289 CGGGGCTTCCGCTGTCGAAGACCATCTTGAATTTTGGGAGGCGTCTTTACGGGTCTCA 3348  
QY 1622 CCCACATAGACGCCCACTTCTTATCCAGAAAGCAGAGTGGGAAACCTTCCCTATC 1681  
DB 3349 CTCATATAGATGCCCACTTCTTATCCAGAAAGCAGAGTGGGAAACCTTCCCTATC 3408  
QY 1682 TGTAGGTTACCAAGCCACGCTGTGCTGCTAGACTCAAGCCCTCCCGCTGTGGGACC 1741  
DB 3409 TGTAGGTTACCAAGCCACGCTGTGCTGCTAGAGCTCAAGCCCTCCCGCTGTGGGACC 3468  
QY 1742 AGATGTGAAGTCTTGTATCCGCTCAAGCCCACTTCCATGGGCAACACCTCTCTAT 1801  
DB 3469 AGATGTGAAGTCTTGTATCCGCTCAAGCCCACTTCCATGGGCAACACCTCTCTAT 3528  
QY 1802 ATAGATGGGCGCTGTCCAGAAATGAAGTCACTTACGACCCAGTCAACAAATATCA 1861  
DB 3529 ACAGACTGGGCGTGTTCAGAAATGAAGTCACTTACGACCCAGTCAACAAATATCA 3588  
QY 1862 TGAATGTATGTGCGGTGACCTGGAGTGTGTCAGAGTACCTGGGTGCTGTGGGGGG 1921  
DB 3589 TGAATGTATGTGCGGTGACCTGGAGTGTGTCAGAGTACCTGGGTGCTGTGGGGGG 3648

QY 1922 TTCTGGCTGCTTTGGCGCGCTATTGCTTATCCACAGGCTGGTGTGTCATAGTAGTAGGA 1981  
DB 3649 TCCTGGCTGCTTTGGCGCGCTATTGCTTATCCACAGGCTGGTGTGTCATAGTAGTAGGA 3708  
QY 1982 TTGTCTTGTCCGAAAGCCGCAATCATACCGACAGGAAAGTCTCTTACCGGAGTTCCG 2041  
DB 3709 TCGTCTTGTCCGAAAGCCGCAATCATACCTGACAGGAAAGTCTCTTACCGAGAGTTCCG 3768  
QY 2042 ATGAATGGAAGAGTCT 2059  
DB 3769 ATGAGATGGAAGAGTCT 3786

## RESULT 9

AAA75296

ID AAA75296 standard; cDNA; 8316 BP.

AC AAA75296;

XX 15-JAN-2001 (first entry)

DT 15-JAN-2001 (first entry)

XX cDNA sequence compiled Hepatitis C virus cDNA clones.

DE Hepatitis C virus; HCV; antisense polynucleotide; polyprotein;

KW viral infectivity; viral replication; ds.

XX Hepatitis C virus.

XX Key Location/Qualifiers

FT CDS 1..8316

FT /\*tag= a

XX /note= "partial sequence; no termination codon given"

PN EP1034785-A2.

XX 13-SEP-2000.

XX 16-MAR-1990; 2000EP-00109602.

XX 17-MAR-1989; 89US-00325338.

PR 20-APR-1989; 89US-00341334.

PR 18-MAY-1989; 89US-00355002.

XX 16-MAR-1990; 90EP-00302866.

XX (CHIR ) CHIRON CORP.

XX Houghton M, Choo Q, Kuo G;

XX WPI; 2000-566891/53.

XX P-PSDB; AAB18540.

XX Novel composition comprising a hepatitis C virus antisense polynucleotide

XX which is complementary to or corresponds to a sense strand of the virus

XX genome, and selectively hybridizes to it.

XX Example; Fig 16; 75pp; English.

XX The specification describes a pharmaceutical composition which comprises

XX a hepatitis C virus (HCV) antisense polynucleotide. The HCV is

XX characterized by a positive stranded RNA genome which has 40% homology at

XX the polypeptide level to a HCV polyprotein. The antisense polynucleotide

XX binds to cellular polynucleotides which enhance and/or are required for

XX viral infectivity, replicative ability or chronicity. The antisense

XX polynucleotides may also be designed to bind with high specificity, to be

XX of increased stability, to be stable and to have low toxicity. The

XX composition also comprises an agent which causes viral RNA to be

XX inactive. The composition is used for preventing HCV replication in a

XX system. The present sequence represents a novel HCV cDNA sequence, which

XX is used in the course of the invention

SQ Sequence 8316 BP; 1671 A; 2529 C; 2345 G; 1771 T; 0 U; 0 Other;

Query Match 86.7%; Score 1786; DB 3; Length 8316;



Best Local Similarity 91.7%; Pred. No. 0; Matches 1888; Conservative 0; Mismatches 170; Indels 0; Gaps 0;

Qy	2	TGGCGCTATCACGGCTATGCCAGCAGACAAGGGCCCTTTTGGGATGCATATACCA	61	Db	3755	CTCTGTCCACCACCGAGAGATCCCTTTTAAOCCGAAGCTATCCCCCTCGAAGTATCA	3814
Db	2735	TGGCGCCATCAACGGCTAGCCACGACAGACAGGGGCCCTCTAGGGTGATATACCA	2794	Qy	1082	AGGGGGGAGAGACATCTCATCTCTGTGCACTCAAAAGAAAGTGCAGACGAGCTCGCGCAA	1141
Qy	62	GCTTGACCGCGCGGACAAAACACAGGTGAGGGTGAAGTTCAGATCGTCAACTGCTG	121	Db	3815	AGGGGGGAGAGACATCTCATCTCTGTGCACTCAAAAGAAAGTGCAGACGAGCTCGCGCAA	3874
Db	2795	GCTTAACTGGCGCGGACAAAACCAAGTGAGGGTGAGTCCAGATTGTGTCAACTGCTG	2854	Qy	1142	AATGTGTGGTGGCGGCTCAATGTCCGTGGCTTATCAACCGCGGCTTGTGTCTCGGTCA	1201
Qy	122	CCGAGACTTTTGGCACTGCAATTAACCGGGTGTGTGGACTGTCTACCATGGAGCGG	181	Db	3875	AGCTGTGCTATTTGGGCACTCAATGTCCGTGGCTTATCAACCGCGGCTTGTGTCTCGGTCA	3934
Db	2855	CCCAAACTTCTTGGCACTGCAATTAACCGGGTGTGTGGACTGTCTACCATGGAGCGG	2914	Qy	1202	TCCGACCAAGTGGTGTGGT	1261
Qy	182	GAAACAAGGACATTTGCGCTACCTTAAGGGTCTGTATCCAGATGTACACCAATGTGGACC	241	Db	3935	TCCGACCAAGTGGTGTGGT	3994
Db	2915	GAACGAGACATTCGCTGTACCCAGGGTCTGTATCCAGATGTACACCAATGTGGACC	2974	Qy	1262	GGGACTTCCATTCGGGTAGTAGTGTCAACACGCTGTGTGTGTGTGTGTGTGTGTGTGTGT	1321
Qy	242	AAGACCTCGTAGGCTGGCCCGCTCCCAAGGTGCGCGCTCATTAACACCATGCACTTGG	301	Db	3995	GGGACTTCCATTCGGGTAGTAGTGTCAACACGCTGTGTGTGTGTGTGTGTGTGTGTGTGT	4054
Db	2975	AAGACCTTGTGGCTGGCCCGCTCCGCAAGGTAGCGCTCATTAACACCATGCACTTGG	3034	Qy	1322	TTGACCTTACCTTCACTTGTAGACAATCAACGCTTCCCGAGGATGTGTCTCCGCTATCTC	1381
Qy	302	GCTCTCGGACCTTACCTGTACGAGGACACGCGATGTATCTGTGTGTGTGTGTGTGTGTGT	361	Db	4055	TTGACCTTACCTTCACTTGTAGACAATCAACGCTTCCCGAGGATGTGTCTCCGCTATCTC	4114
Db	3035	GCTCTCGGACCTTACCTGTACGAGGACACGCGATGTATCTGTGTGTGTGTGTGTGTGTGT	3094	Qy	1382	AACGTCGGGTAGGACTGTGACAGGGAAGCCAGGCACTACAGATTGTGGCAACCGGGG	1441
Qy	362	GTGATGACGGGACGCTGCTGTGCGCCCGGCCATTTCTCTACTTGAAGGCTCTCTCGG	421	Db	4115	AACGTCGGGTAGGACTGTGACAGGGAAGCCAGGCACTACAGATTGTGGCAACCGGGG	4174
Db	3095	GTGATGACGGGACGCTGCTGTGCGCCCGGCCATTTCTCTACTTGAAGGCTCTCTCGG	3154	Qy	1442	AGCGCTCTTCTGCGATGTTTGTACTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCT	1501
Qy	422	GAGGCCCTCTGCTGTGCGCGCAGGACATGCGGTAGGCATATTCAGAGCCCGCGTATGCA	481	Db	4175	AGCGCTCTTCTGCGATGTTTGTACTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCT	4234
Db	3155	GGGTCGCTGTGTGCGCGCGCGGCAACGCGGTGGGCATATTTAGGGCCGCGGTGTGCA	3214	Qy	1502	CTTGTATGAGCTTACGCGCGCGAGACCAACAGCTTACAGATTGTGGCAACCGGCC	1561
Qy	482	CCGCTGAGTGTGTAAGCGGGTGAATTCATTCGCGGTAGAGCTTATAGACAAACATGA	541	Db	4235	CTTGTATGAGCTTACGCGCGCGAGACTACAGTTAGGCTACGAGCTACATGAACACCC	4294
Db	3215	CCGCTGAGTGTGTAAGCGGGTGAATTCATTCGCGGTAGAGCTTATAGACAAACATGA	3274	Qy	1562	CGGGACTTCCGCTGTGCGCAAGACCATCTTGAATTTTGGGAGGGCGTCTTTACGGGTCTCA	1621
Qy	542	GCTTCCCGGTGTTCTCAGACAATCTCTCCACACAGAGTCCCGGCGGCTACCAAGTGG	601	Db	4295	CGGGGCTTCCGCTGTGCGCAAGACCATCTTGAATTTTGGGAGGGCGTCTTTTACAGGCTCA	4354
Db	3275	GGTCCCGGTGTTCTCAGACAATCTCTCCACAGAGTCCCGGCGGCTTCCAGGTGG	3334	Qy	1622	CCACATAGACGCGCACTCTCTATCCAGACAAAGCAGAGTGGGGAACCTTCCCTATC	1681
Qy	602	CCACCTGCTATGCTCCACCGGACGCGTAAAGACACCAAGGTCCCGGCGGCTACGCGAG	661	Db	4355	CTCATATAGATGCCCACTTCTATCCAGACAAAGCAGAGTGGGGAACCTTCTCTTACC	4414
Db	3335	CTCACCTCATGCTCCACAGGACGCGGAAAGACCAAGGTCCCGGCTGCATATGCGAG	3394	Qy	1682	TGTAGCTGACCAAGCACCGCTGTGCGCTAGAGCTCAAGCCCCCTCCCGCTGTGGGACC	1741
Qy	662	CTCAGGGCTACAAGGTGCTGGTGTCAACCCCTCGGTGTGTGCAACATGGGCTTTGGTG	721	Db	4415	TGTAGCTGACCAAGCACCGCTGTGCGCTAGAGCTCAAGCCCCCTCCCGCTGTGGGACC	4474
Db	3395	CTCAGGGCTACAAGGTGCTGGTGTCAACCCCTCGGTGTGTGCAACATGGGCTTTGGTG	3454	Qy	1742	AGATGTGGAAGT	1801
Qy	722	CTTACATGTCCAAGGCCATGGGATTCATCTTAACATCAGGACTGGGGTGAGACAATTA	781	Db	4475	AGATGTGGAAGT	4534
Db	3455	CTTACATGTCCAAGGCCATGGGATTCATCTTAACATCAGGACTGGGGTGAGACAATTA	3514	Qy	1802	ATAGACTGGGCGCTGTCCAGAAATGAAGTCAACCTGACGACACCCAGTCAACCAAGTATATCA	1861
Qy	782	CTTGTGCGCCGCTATGACATAAATTTGTGACGAGTGCCTACCGCAAGTTCCTTTCGCGAGCGGGGTGT	841	Db	4535	ACAGACTGGGCGCTGTTCAGAAATGAATCAACCTGACGACCCAGTCAACCAATATATCA	4594
Db	3515	CCACTGCGCCGCTATGACATAAATTTGTGACGAGTGCCTACCGCAAGTTCCTTTCGCGAGCGGGGTGT	3574	Qy	1862	TGACATGTATGTGGGCTGACTGTGAGGTGCTCAACGATACCTGGGTGCTCGTTTGGCGGG	1921
Qy	842	CAGGGGTGCTTATGACATAAATTTGTGACGAGTGCCTACCGCAAGTTCCTTTCGCGAGCGGGGTGT	901	Db	4595	TGACATGTATGTGGGCTGACTGTGAGGTGCTCAACGATACCTGGGTGCTCGTTTGGCGGG	4654
Db	3575	CGGGGGCGCTTATGACATAAATTTGTGACGAGTGCCTACCGCAAGTTCCTTTCGCGAGCGGGGTGT	3634	Qy	1922	TTCTGGCTCTTGGCGCGTATTCCTATCCAGGGCTGCGGTGCTCATAGTAGTAGGA	1981
Qy	902	TCTTGGCATTTGGCACTGCTTGACCAAGCAGAGACCGGGGGGCGAGACTGACTGTGC	961	Db	4655	TCTGGCTCTTGGCGCGTATTCCTATCCAGGGCTGCGGTGCTCATAGTAGTAGGA	4714
Db	3635	TCTTGGCATTTGGCACTGCTTGACCAAGCAGAGACTGCGGGGGGCGAGACTGCTGTGTGC	3694	Qy	1982	TTGTCTTGTCCGGAAGCGGCAATCATACCCGACAGGGAAGTCTCTCTACCGGGAGTTCG	2041
Qy	962	TGCGCACCGCTACCCCTCCGGGTCTCGTCACTGTGCCCATCTCTAAATCGAGAGGTG	1021	Db	4715	TGCTCTTGTCCGGAAGCGGCAATCATACCTGACAGGGAAGTCTCTCTACCGAGAGTTCG	4774
Db	3695	TGCGCACCGCTACCCCTCCGGGTCTCGTCACTGTGCCCATCTCTAAATCGAGAGGTG	3754	Qy	2042	ATGAAATGGAAGAGTGTCT 2059	
Qy	1022	CTCTGTCCACTACCGGAGAGATCCCTTTTATGTGCAAGGCTATTTCCCTTTGAAGCAATTA	1081	Db	4775	ATGAGATGGAAGAGTGTCT 4792	

RESULT 10  
AA207656



AAZ07656 standard; DNA; 9133 BP.  
AAZ07656;  
20-MAR-2003 (revised)  
08-NOV-1999 (first entry)  
Nucleotide sequence of HCV-1 ORF.  
Hepatitis C virus; HCV; J1; J7; HCV-1; non-A, non-B HCV; NANBH;  
HCV infection; vaccine; ds.  
Hepatitis C virus.  
OS Location/Qualifiers  
FH CDS  
FT 288..9132  
FT /\*tag= b  
FT /transl\_except= (pos:1588..1589; aa:Leu)  
FT /note= "this codon has an apparent 1 nucleotide deletion,  
FT which alters the reading frame"  
FT /transl\_except= (pos:1647..1650; aa:Pro)  
FT /note= "this codon has an apparent 1 nucleotide  
FT insertion, which alters the reading frame; this insertion  
FT is not indicated in the sequence present in the formal  
FT sequence listing of the specification"  
XX  
EP939128-A2.  
XX  
XX  
XX 01-SEP-1999.  
XX 17-SEP-1990; 99EP-00101746.  
XX  
XX 15-SEP-1989; 89US-00408045.  
XX 21-DEC-1989; 89US-00456142.  
XX 17-SEP-1990; 90EP-00310149.  
XX  
XX (OYAA/) OYA A.  
XX (CHIR ) CHIRON CORP.  
XX  
XX Miyamura T, Saito I, Houghton M, Weiner AJ, Han J, Kolberg JA;  
XX Cha T, Irvine BD;  
XX P-PSDB; AAV14975.  
XX  
XX WPI; 1999-480843/41.  
XX  
XX New Hepatitis C Virus isolates, useful for diagnosis of hepatitis  
XX infections and development of vaccines.  
XX  
XX Disclosure; Fig 12; 132pb; English.  
XX  
XX The invention provides two new isolates of hepatitis C virus (HCV), J1  
XX and J7. These two isolates comprise nucleotide and amino acid sequences  
XX that are distinct from the HCV isolate HCV-1. The nucleotide sequences  
XX may be used to detect non-A, non-B HCV (NANBH) polynucleotides by  
XX hybridization for diagnosis of NANBH infections. They may also be used to  
XX screen blood donors, donated blood and blood products for this infection.  
XX The isolates may also be used to isolate other naturally occurring  
XX variants of the virus. The polypeptides may be used as a vaccine for  
XX administration to patients to protect against infection with NANBH. The  
XX present sequence represents the nucleotide sequence of HCV-1 ORF.  
XX (Updated on 20-MAR-2003 to correct PF field.) (Updated on 20-MAR-2003 to  
XX correct PR field.)  
XX  
XX Sequence 9133 BP; 1834 A; 2772 C; 2600 G; 1927 T; 0 U; 0 Other;  
XX  
XX Query Match 86.7%; Score 1786; DB 2; Length 9133;  
XX Best Local Similarity 91.7%; Pred. No. 0;  
XX Matches 1888; Conservative 0; Mismatches 170; Indels 0; Gaps 0;  
XX  
XX 2 TGGCGCCTATCGCGGCTATGCCAGCAGACAGAGGGGCTTTGGGATGCATATCACCA 61  
XX |||||||  
XX 3344 TGGCGCCCATCACGGGCTACGCGGCTAGCCCAAGAGGGGCTCTTAGGGTGATATCACCA 3403  
XX |||||||



482 CCGTGGAGTGGCTAAGCGGTGGACTTCATCCCGTAGAGAGCTTTAGAGACAACCATCA 541  
Db |||||  
3876 CCGTGGAGTGGCTAAGCGGTGGACTTCATCCGTGGAGAACCTTAGAGACAACCATCA 3935  
QY |||||  
542 GGTCCCGGTGTTCTCAGACAACCTCTCCACAGCAGTGTGCCAGAGCTTCCAAAGTGG 601  
Db |||||  
3936 GGTCCCGGTGTTCTCAGGATAAATCTCTCCACAGTAGTGTGCCAGAGCTTCCAGGTGG 3995  
QY |||||  
602 CCCACCTGCATGCTCCACCGGAGCGGTAAAGAGCACCAAGGTCCCGGCGCATACGCAG 661  
Db |||||  
3996 CTCACCTCCATGCTCCACAGGAGCGGCAAAAGCACCAAGGTCCCGGCTGCATATGCAG 4055  
QY |||||  
662 CTCAGGGCTACAGGTGCTGCTCAACCCCTCTCGTTGCTGCAACAATGGGCTTTGGTG 721  
Db |||||  
4056 CTCAGGGCTATAAGGTGCTAGTACTCAACCCCTCTGTTGCTGCAACACCTGGGCTTTGGTG 4115  
QY |||||  
722 CTTACATGTCGAAGGCCATGGGATTGATCTTAACATCAGAGCTGGGGTGAGGACAATTA 781  
Db |||||  
4116 CTTACATGTCGAAGGCTCATGGATCGATCCTTAACATCAGGACCGGGGTGAGAACAAATTA 4175  
QY |||||  
782 TACTGGCAGCCGATCAGTATTTCCACCTACGGCAAGTTCTTGGCCGACGCGGTGT 841  
Db |||||  
4176 CCACTGGCAGCCCAATCAGTACTCCACCTACGGCAAGTTCTTGGCCGACGCGGTGT 4235  
QY |||||  
842 CAGGGGTGCTTATGACATAATAATTTGTGACAGGTGCCACTCCACGATGCAACATCCA 901  
Db |||||  
4236 CGGGGGCGCTTATGACATAATAATTTGTGACAGGTGCCACTCCACGATGCAACATCCA 4295  
QY |||||  
902 TCTTGGGCAATGGCACTGTCTTGACAAAGCAGAGACCGCGGGGGGAGACTGACTGTGC 961  
Db |||||  
4296 TCTTGGGCAATCGGCACTGTCTTGACAAAGCAGAGACTCGGGGGGAGACTGGTTGTGC 4355  
QY |||||  
962 TGGCCACCGTACCCCTCGGGCTCGTCACTGTGCCCATCTTAACATCGAGGAGTTG 1021  
Db |||||  
4356 TGGCCACCGTACCCCTCGGGCTCGTCACTGTGCCCATCTTAACATCGAGGAGTTG 4415  
QY |||||  
1022 CTCTGTCCACTACCGGAGAGATCCCTTTTATGGCAAGGCTATTCCTTTGAAGCAATTA 1081  
Db |||||  
4416 CTCTGTCCACTACCGGAGAGATCCCTTTTATGGCAAGGCTATTCCTTTGAAGCAATTA 4475  
QY |||||  
1082 AGGGGGGAGACATCTCATCTTCTGCACTCAAGAGAGTGCAGAGCTCGCGGAA 1141  
Db |||||  
4476 AGGGGGGAGACATCTCATCTTCTGCACTCAAGAGAGTGCAGAGCTCGCGGAA 4535  
QY |||||  
1142 AACTGTGTGCTTGGCGCTCAATGCGTGTCTTACTACCGCGCTTGTGATGTGCTGCA 1201  
Db |||||  
4536 AGCTGTGTGCTTGGCGCTCAATGCGTGTCTTACTACCGCGCTTGTGATGTGCTGCA 4595  
QY |||||  
1202 TCCGACCAAGTGTGACGTGTGCTGCTGGCAACTGACCGCTCTCATGACCGGCTTTACCG 1261  
Db |||||  
4596 TCCGACCAAGTGTGACGTGTGCTGCTGGCAACTGACCGCTCTCATGACCGGCTTTACCG 4655  
QY |||||  
1262 GCGACTTCGATTCGGGTAGACTGCAACGCTGTGTCTACCCAGACAGTGCAGCTTCAGCC 1321  
Db |||||  
4656 GCGACTTCGATTCGGGTAGACTGCAACGCTGTGTCTACCCAGACAGTGCAGCTTCAGCC 4715  
QY |||||  
1322 TTGACCTTACCTTCCACTTGAACAATCACTGCTTCCCGAGATGTGTCTCCCGTACTC 1381  
Db |||||  
4716 TTGACCTTACCTTCCACTTGAACAATCACTGCTTCCCGAGATGTGTCTCCCGTACTC 4775  
QY |||||  
1382 AACGTCCGGGTAGGACTGGCAGAGGAAGCCAGGCAATCTACAGATTTGTGGCACCGGGG 1441  
Db |||||  
4776 AACGTCCGGGTAGGACTGGCAGAGGAAGCCAGGCAATCTACAGATTTGTGGCACCGGGG 4835  
QY |||||  
1442 AGGCTCTTCTGGCATGTTGACTCGTCTGTCTCTGCGAGTGTCTATGACGGGGTGTG 1501  
Db |||||  
4836 AGGCTCTTCTGGCATGTTGACTCGTCTGTCTCTGCGAGTGTCTATGACGGGGTGTG 4895  
QY |||||  
1502 CTTGGTATGAGCTTACCGCGCGGAGACCACTAGTGTAGGCTACGAGATACATGAACACC 1561  
Db |||||  
4896 CTTGGTATGAGCTTACCGCGCGGAGACCACTAGTGTAGGCTACGAGGCTACATGAACACC 4955  
QY |||||  
1562 CGGGACTTCCCGTGTGCCAAGACCATCTTGAATTTTGGGAGGGCGTCTTTACGGGTCTCA 1621

4956 CGGGCTTCCCGTGTGCCAGGACCATCTGAATTTTGGAGGCGCTTTTACAGGCTCA 5015  
QY |||||  
1622 CCCACATAGACGCCCACTTCTATCCAGACAAGCAGAGTGGGGAACCTTCCCTATC 1681  
Db |||||  
5016 CTCATATAGATGCCCACTTCTATCCAGACAAGCAGAGTGGGGAACCTTCCCTATC 5075  
QY |||||  
1682 TGGTAGCGTACCAAGCCACCGTGTGCTAGAGCTCAAGCCCTCCCGTGTGGGACC 1741  
Db |||||  
5076 TGGTAGCGTACCAAGCCACCGTGTGCTAGAGCTCAAGCCCTCCCGTGTGGGACC 5135  
QY |||||  
1742 AGATGTGGAAGTGTGATTCGTCTCAAGCCACCTCATGGGCCAACACCTCTGCTAT 1801  
Db |||||  
5136 AGATGTGGAAGTGTGATTCGTCTCAAGCCACCTCATGGGCCAACACCTCTGCTAT 5195  
QY |||||  
1802 ATAGACTGGGCGCTGTCCAGATGAAGTCAACCTCAGCACCAGTCAACCAAGTATATCA 1861  
Db |||||  
5196 ACAGACTGGGCGCTGTTCAGAAATGAATCAACCTCAGCACCAGTCAACCAATATATCA 5255  
QY |||||  
1862 TGACATGTATGTGGCTGACCTGGAGGTCTGTCAGAGTACCTGGGTGCTCTGTTGGCGG 1921  
Db |||||  
5256 TGACATGTATGTGGCGGACCTGGAGGTCTGTCAGAGTACCTGGGTGCTCTGTTGGCGG 5315  
QY |||||  
1922 TTCTGGCTGTGTTGGCGGCTATTCCTATCCAGAGTGTGCTCATAGTAGTAGGA 1981  
Db |||||  
5316 TCCTGGCTGTGTTGGCGGCTATTCCTATCCAGAGTGTGCTCATAGTAGTAGGA 5375  
QY |||||  
1982 TTGCTTGTGCGGAAGCGGCAATCATCCGACAGGAGTCTCTACCGGAGTTCG 2041  
Db |||||  
5376 TCCTTGTGCGGAAGCGGCAATCATCCGACAGGAGTCTCTACCGGAGTTCG 5435  
QY |||||  
2042 ATGAATGGAAGTGTCT 2059  
Db |||||  
5436 ATGAGTGAAGAGTGTCT 5453

RESULT 12  
AAA75297  
ID AAA75297 standard; cdna; 9185 BP.  
XX AC AAA75297;  
XX DT 15-JAN-2001 (first entry)  
XX DE Sense strand of HCV encoding a polyprotein.  
XX KW Hepatitis C virus; HCV; antisense polynucleotide; polyprotein;  
XX OS viral infectivity; viral replication; ds.  
XX OS Hepatitis C virus.  
XX FH Key  
XX CDS Location/Qualifiers  
FT 320..9184  
FT /\*tag= a  
FT /note= "partial sequence; no termination codon given"  
XX  
XX EP1034785-A2.  
XX  
XX PD 13-SEP-2000.  
XX  
XX PF 16-MAR-1990; 2000EP-00109602.  
XX  
XX PR 17-MAR-1989; 89US-00325338.  
XX PR 20-APR-1989; 89US-00341334.  
XX PR 18-MAY-1989; 89US-00355002.  
XX PR 16-MAR-1990; 90EP-00302866.  
XX  
XX PA (CHIR ) CHIRON CORP.  
XX  
XX Houghton M, Choo Q, Kuo G;  
XX WPI; 2000-566891/53.  
XX P-PSDB; AAB18541.  
DR

XX Novel composition comprising a hepatitis C virus antisense polynucleotide  
PT which is complementary to or corresponds to a sense strand of the virus  
PT genome, and selectively hybridizes to it.  
XX Example; Fig 17; 75pp; English.

XX The specification describes a pharmaceutical composition which comprises  
CC a hepatitis C virus (HCV) antisense polynucleotide. The HCV is  
CC characterized by a positive stranded RNA genome which has 40% homology at  
CC the polypeptide level to a HCV polypeptide. The antisense polynucleotide  
CC binds to cellular polynucleotides which enhance and/or are required for  
CC viral infectivity, replicative ability or chronicity. The antisense  
CC polynucleotides may also be designed to bind with high specificity, to be  
CC of increased stability, to be stable and to have low toxicity. The  
CC composition also comprises an agent which causes viral RNA to be  
CC inactive. The composition is used for preventing HCV replication in a  
CC system. The present sequence represents a novel HCV cDNA sequence, which  
CC is used in the course of the invention

XX Sequence 9185 BP; 1849 A; 2790 C; 2608 G; 1938 T; 0 U; 0 Other;

Query Match 86.7%; Score 1786; DB 3; Length 9185;

Best Local Similarity 91.7%; Pred. No. 0;

Matches 188; Conservative 0; Mismatches 170; Indels 0; Gaps 0;

QY	2	TGGCGCCCTATCAGCGCCCTATGCCAGCAGACAAAGGGGCCCTTTGGGATGCATATCACCA	61
DB	3396	TGAGGCCCATCAGCGCGTAGCCACGACAGACAGAGGGCCCTCTAGGGTGATATCACCA	3455
QY	62	GCTTGACCGCCCGGACAAAACACAGTGAGGGTGAGGTTAGATGCTCAACTGCTG	121
DB	3456	GCCTAACTGCGCGGACAAAACCAAGTGAGGGTGAGGTTAGATGCTCAACTGCTG	3515
QY	122	CCGAGCTTTCTGGCAACTGCATTAACGGGGTGTTGGAGTGTCTACCATGGAGCCG	181
DB	3516	CCCAACCTTCTGGCAACTGCATTAACGGGGTGTTGGAGTGTCTACCATGGAGCCG	3575
QY	182	GAACAAGGACATTCGCTCACCTTAAGGGTCTGTATCCAGATGTACCAATGTGGACC	241
DB	3576	GAAGGAGGACATTCGCTCACCCNAGGGTCTGTATCCAGATGTATCCNATGTAGACC	3635
QY	242	AAGACTCGTAGGTCGCGCTCCCAAGGTGCGCGCTCATTAACCAATGCATTTGG	301
DB	3636	AAGACTTGTGGCTGCGCGCTCCGCAAGGTAGCGCTCATTAACCAATGCATTTGG	3695
QY	302	GCTTCTCGGACCTTACCTGGTCAAGGACACCGCGATGTCTCTGTGCGCGGACGG	361
DB	3696	GCTTCTCGGACCTTACCTGGTCAAGGACACCGCGATGTCTCTGTGCGCGGACGG	3755
QY	362	GTGATGGAGGGGAGCTCTTTTGGCCCGGCTATCTCTTACTTGAAGGCTCTCTCGG	421
DB	3756	GTGATGGAGGGGAGCTCTGTGCGCCCGGCGCATTTCTTACTTGAAGGCTCTCTCGG	3815
QY	422	GAGGCCCTCTGTGTGCGCCGAGGACATGCCGTAGGCATATTACAGCGCGGATGCA	481
DB	3816	GAGGCTCGGTGTGTGCGCCGCGGACGCGCTGGGACATATTAGGGCGCGGTGTGCA	3875
QY	482	CCGCTGGAGTGGCTAAGCGGCTGATTTATCCCGGTAGAGACTTAGAGACAACATGA	541
DB	3876	CCGCTGGAGTGGCTAAGCGGCTGATTTATCCCGGTAGAGACTTAGAGACAACATGA	3935
QY	542	GGTCCCGGTGTCTCAGACAACTCTTCCCGACAGCAGTCCCGAGAGCTACCAAGTGG	601
DB	3936	GGTCCCGGTGTCTCAGAGTAATCTCTCTCCAGTAGTGCCCGAGAGCTTCCAGGTGG	3995
QY	602	CCCACTGCAATGCTCCACCGGAGGGTAAAGAGCACCAAGGTCCCGGCGCATAGCAG	661
DB	3996	CTCACTCCATGCTCCACAGGAGCGGCAAAAGACCAAGGTCCCGGCTGCATATGCAG	4055
QY	662	CTCAGGGCTACAAGGTGCTGCTCAACCCCTCCGTTGCTGCAACATGGGCTTTGGTG	721
DB	4056	CTCAGGGCTACAAGGTGCTGCTCAACCCCTCTGTTGCTGCAACACTGGGCTTTGGTG	4115

QY	722	CTTACATGTCCAAGGCCCATGGGATGTATCTTAACATCAGGACTGGGGTGAGACAATTA	781
DB	4116	CTTACATGTCCAAGGCTCATGGATGATCTTAACATCAGGACCGGGTGAGACAATTA	4175
QY	782	CTACTGGAGCGGATCATCGTATTCACCTACGGCAAGTTCCTTGGCGAGCGGGTGT	841
DB	4176	CCATCTGGACCGCCCATCAGCTACTCCACCTACGGCAAGTTCCTTGGCGAGCGGGTGT	4235
QY	842	CAGGGGGTGTATGACATAAATTTGTGACAGGTGCCACTCCACGAGTGAACATCCA	901
DB	4236	CGGGGGGCGTATGACATAAATTTGTGACAGGTGCCACTCCACGAGTGAACATCCA	4295
QY	902	TCTTGGGCATTTGGCACTGTCTTGTACCAAGCAGAGACCGCGGGGGGAGACTGTGTC	961
DB	4296	TCTTGGGCATTTGGCACTGTCTTGTACCAAGCAGAGACTCGGGGGGAGACTGTGTC	4355
QY	962	TGGCCACCGCTACCCCTCGGGGTCGGTCACTGTGTCGCCCATCTTAACATCGAGGAGT	1021
DB	4356	TGGCCACCGCCACCCCTCGGGGTCGGTCACTGTGTCGCCCATCTTAACATCGAGGAGT	4415
QY	1022	CTCTGTCCACTACCGGAGAGATCCCTTTTATGGCAAGCTATTCCTTGAAGCAATTA	1081
DB	4416	CTCTGTCCACTACCGGAGAGATCCCTTTTATGGCAAGCTATTCCTTGAAGCAATTA	4475
QY	1082	AGGGGGGAGACATCTCATCTTCTGCACTCAAAAGAGAGTGCAGAGCTCGCGCAA	1141
DB	4476	AGGGGGGAGACATCTCATCTTCTGCACTCAAAAGAGAGTGCAGAGCTCGCGCAA	4535
QY	1142	AACTGTGCGTGGGCGTCAATGCGTGGCTTATCTACCGGGCTTGTATGTGTCGTCA	1201
DB	4536	AGCTGTGCGTGGGCGTCAATGCGTGGCTTATCTACCGGGCTTGTATGTGTCGTCA	4595
QY	1202	TCCGACAGTGTGAGTGTGCTGTCGCAACTGACCGCCCTCATGACCGGCTTTACCG	1261
DB	4596	TCCGACAGTGTGAGTGTGCTGTCGCAACTGACCGCCCTCATGACCGGCTTTACCG	4655
QY	1262	GCGACTTCCATTCGGGTGATAGACTGCAACACGCTGTCTACCCAGAGAGTGCAGCT	1321
DB	4656	GCGACTTCCATTCGGGTGATAGACTGCAACACGCTGTCTACCCAGAGAGTGCAGCT	4715
QY	1322	TTGACCTTACCTTACCATTTAGAGCAATCACTGCTTCCCGAGATGTCTTCCCGTCTC	1381
DB	4716	TTGACCTTACCTTACCATTTAGAGCAATCACTGCTTCCCGAGATGTCTTCCCGTCTC	4775
QY	1382	AACTGTGCGGTAGGACTGGCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	1441
DB	4776	AACTGTGCGGTAGGACTGGCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	4835
QY	1442	AGCGTCTTCTGCGCATGTTGACTGCTGCTCTCTGCGAGTGTCTATGACCGGGTGTG	1501
DB	4836	AGCGTCTTCTGCGCATGTTGACTGCTGCTCTCTGCGAGTGTCTATGACCGGGTGTG	4895
QY	1502	CTTGTGTAGCTTACGCGCGCGAGACCACTAGTGTAGGCTACGAGCATCATGAAACCC	1561
DB	4896	CTTGTGTAGCTTACGCGCGCGAGACTACAGTGTAGGCTACGAGCATCATGAAACCC	4955
QY	1562	CGGGACTTCCCGTGTGCGCAAGACCATCTTGAATTTTGGAGGGGCGTCTTACGGGTCTCA	1621
DB	4956	CGGGACTTCCCGTGTGCGCAAGACCATCTTGAATTTTGGAGGGGCGTCTTACGGGTCTCA	5015
QY	1622	CCACATAGACGCCCACTCTCTATCCAGACAAAGCAGAGTGGGAGGAGGAGGAGGAGG	1681
DB	5016	CTCATATAGTGCCTCACTTCTATCCAGACAAAGCAGAGTGGGAGGAGGAGGAGGAGG	5075
QY	1682	TGTAGCTTACCAAGCCACCGTGTGCGTGTAGAGTCAAGCCCTTCCCGGTGTGGAGC	1741
DB	5076	TGTAGCTTACCAAGCCACCGTGTGCGTGTAGAGTCAAGCCCTTCCCGGTGTGGAGC	5135
QY	1742	AGATGTGGAAGTGTGTGCTTCAAGCCACCTCTCAAGCCACCTCTCAAGCCACCTCTCTAT	1801
DB	5136	AGATGTGGAAGTGTGTGCTTCAAGCCACCTCTCAAGCCACCTCTCAAGCCACCTCTCTAT	5195

QY 1802 ATAGACTGGCGCTGTCAGAAATGAAGTCACCTGACGACCCAGTCACCAAGTATATCA 1861  
DB 5196 ACAGACTGGCGCTGTTCAAGATGAATCACCCTGACGACCCAGTCACCAATATATCA 5255  
QY 1862 TGACATGTATGTCGGCTGACCTGGAGGTGCTCAGAGTACCTGGGTGCTCGTTGGCGGG 1921  
DB 5256 TGACATGTATGTCGGCTGACCTGGAGGTGCTCAGAGTACCTGGGTGCTCGTTGGCGGG 5315  
QY 1922 TTCTGGCTGCTTTGGCGCGGTATGCTCAATGATGCTCAACAGGCTGCTCATAGTAGTAGGA 1981  
DB 5316 TCCTGGCTGCTTTGGCGCGGTATGCTCAACAGGCTGCTCATAGTAGTAGGA 5375  
QY 1982 TTGCTTTGTCGGAAGCGGCAATCATACCCAGAGGAAGTCTCTACCGGAGTTGG 2041  
DB 5376 TGTCTTTGTCGGAAGCGGCAATCATACCTGACAGGGAAGTCTCTACCGAGAGTTGG 5435  
QY 2042 ATCAAAATGAAGAGTGCT 2059  
DB 5436 ATGAGATGAAGAGTGCT 5453

## RESULT 13

ADN35979  
ID ADN35979 standard; cDNA; 9185 BP.

AC ADN35979;

XX 17-JUN-2004 (first entry)

DT HCV cDNA clone #2.

DE Antiviral; Vaccine; hepatitis C virus infection; HCV infection; ss.

KW Antiviral; Vaccine; hepatitis C virus infection; HCV infection; ss.

XX Hepatitis C virus.

OS BP1394255-A2.

PN 03-MAR-2004.

PD 16-MAR-1990; 2003EP-00016585.

PF 17-MAR-1989; 89US-00325338.

PR 20-APR-1989; 89US-00341334.

PR 18-MAY-1989; 89US-00355002.

PR 16-MAR-1990; 90EP-00302866.

XX (CHIR ) CHIRON CORP.

PA Houghton M, ; Choo Q, Kuo G;

XX WPI; 2004-193149/19.

DR P-PSDB; ADN35978.

DR Novel purified hepatitis C virus polypeptide comprising epitope encoded

XX by hepatitis C virus cDNA, useful as vaccine for treating hepatitis C

PT virus.

PT Claim 1; Fig 17; 79pp; English.

XX The present invention relates to hepatitis C virus (HCV) proteins and

CC cDNA sequences. The sequences are useful in immunoassays for detecting

CC antibodies directed against HCV antigen; preparing host cells transformed

CC with a recombinant polynucleotide; screening antiviral agents and

CC determining the effect of antiviral agent in inhibiting viral replication

CC in cell culture system; and developing vaccine for treating HCV

CC infection.

XX Sequence 9185 BP; 1849 A; 2790 C; 2608 G; 1938 T; 0 U; 0 Other;

QY Query Match 86.7%; Score 1786; DB 12; Length 9185;

DB Best Local Similarity 91.7%; Pred. No. 0;

Matches 1888; Conservative 0; Mismatches 170; Indels 0; Gaps 0;

QY 2 TGGCGCTTATCACGCGCTATGCCAGCAGACAAGGGGCTTTTGGATGCAATATACCA 61  
DB 3396 TGGCGCTTATCACGCGCTATGCCAGCAGACAAGGGGCTTTTGGATGCAATATACCA 3455  
QY 62 GCTTGACCGGCGGGACAAAAACAGGTGGAGGTGAGGTTGAGATCGTGTCAACTGTG 121  
DB 3456 GCTTAACCTGGCGGGACAAAAACAGGTGGAGGTGAGGTTGAGATCGTGTCAACTGTG 3515  
QY 122 CCCAGACTTTCTTGGCAACCTGCAATTAACGGGCTGTGTGGACTGTCTACCATGGAGCG 181  
DB 3516 CCCAAACCTTCTTGGCAACCTGCAATTAACGGGCTGTGTGGACTGTCTACCATGGAGCG 3575  
QY 182 GAAACAAGACCAATTCGCTCAGCTTAAGGCTCTGTTATCCAGATGTACCAATGTGGACC 241  
DB 3576 GAAACAAGACCAATTCGCTCAGCTTAAGGCTCTGTTATCCAGATGTACCAATGTGGACC 3635  
QY 242 AAGACTCTGTAGGCTGGCGCGCTCCCAAGGTGCCGCTCATTAACATGCACTTGGG 301  
DB 3636 AAGACTCTGTAGGCTGGCGCGCTCCCAAGGTGCCGCTCATTAACATGCACTTGGG 3695  
QY 302 GCTCCTCGGACCTTTTACCTGCTCAGAGGACGCGGATGTATTCTCTGTGGCGGACGCG 361  
DB 3696 GCTCCTCGGACCTTTTACCTGCTCAGAGGACGCGGATGTATTCTCTGTGGCGGACGCG 3755  
QY 362 GTGATGGCAGGGGAGCCTGCTTTTGGCGCGGCTTATCTCTTACTTGAAGGCTCTCGG 421  
DB 3756 GTGATGGCAGGGGAGCCTGCTTTTGGCGCGGCTTATCTCTTACTTGAAGGCTCTCGG 3815  
QY 422 GAGGCCCTGCTGTGGCGCGGACGATGCGGTAGGATATTACAGCGCGGCTATGCA 481  
DB 3816 GAGGCCCTGCTGTGGCGCGGACGATGCGGTAGGATATTACAGCGCGGCTATGCA 3875  
QY 482 CCGCTGGAGTGGCTAAGCGGTGGACTTTCATCCCGTAGAGAGCTTAGAGACAACCATGA 541  
DB 3876 CCGCTGGAGTGGCTAAGCGGTGGACTTTCATCCCGTAGAGAGCTTAGAGACAACCATGA 3935  
QY 542 GGTCCCGGCTGTCTCAGACAACTCCTCCCAACAGCAGTGCGCCAGAGCTTACCAAGTGG 601  
DB 3936 GGTCCCGGCTGTCTCAGACAACTCCTCCCAACAGCAGTGCGCCAGAGCTTACCAAGTGG 3995  
QY 602 CCCACTGCTATGCTCCGCGGACGCGGTAGAGCACCAGGTCGCCGCGCATACGAG 661  
DB 3996 CTCACCTCATGCTCCGCGGACGCGGTAGAGCACCAGGTCGCCGCGCATACGAG 4055  
QY 662 CTCAGGGCTTACAAGGTGCTGTGCTCAACCCCTCCGCTGCTGTCGAACAATGGCTTTGGTG 721  
DB 4056 CTCAGGGCTTACAAGGTGCTGTGCTCAACCCCTCCGCTGCTGTCGAACAATGGCTTTGGTG 4115  
QY 722 CTTACATGTCCAAGGCCCATGGATTGATCTTAAACATCAGGAGTGGGTGAGGACAATTA 781  
DB 4116 CTTACATGTCCAAGGCCCATGGATTGATCTTAAACATCAGGAGTGGGTGAGGACAATTA 4175  
QY 782 CTACTGGCAGCCCGATCAGTATTCACCTAGCGAAGTTCCTTCCGACGCGGGTGT 841  
DB 4176 CCACTGGCAGCCCGATCAGTATTCACCTAGCGAAGTTCCTTCCGACGCGGGTGT 4235  
QY 842 CAGGGGTGCTTATGACATAATAATTTGTGAGAGTGGCCACTTCCACGGATGCAACATCCA 901  
DB 4236 CAGGGGTGCTTATGACATAATAATTTGTGAGAGTGGCCACTTCCACGGATGCAACATCCA 4295  
QY 902 TCTTGGGCAATGGCATGCTGCTTGAACAGAGAGACCGGGGGGAGAGCTGACTGTGC 961  
DB 4296 TCTTGGGCAATGGCATGCTGCTTGAACAGAGAGACTGCGGGGGGAGAGCTGCTGTGC 4355  
QY 962 TCGCCACCGCTACCCCTCGGCTCGCTCAGTGTGCCCATCTTAACATCGAGGAGGTG 1021  
DB 4356 TCGCCACCGCTACCCCTCGGCTCGCTCAGTGTGCCCATCTTAACATCGAGGAGGTG 4415  
QY 1022 CTCTGTCCACTACCGGAGAGATCCCTTTTATGGCAAGGCTATTCCCTTGAAGCAATTA 1081  
DB 4416 CTCTGTCCACTACCGGAGAGATCCCTTTTATGGCAAGGCTATTCCCTTGAAGCAATTA 4475  
QY 1082 AGGGGGGGAGACATCTCATCTTCTGCCACTCAAGAGAGAGTGGCAGAGAGCTCGCGCAA 1141

Db 4476 AGGGGGGAGACATCTCATCTTCTGTCATTCAGAGAGAAAGTGGCAGAACTCGCGCAA 4535  
Qy 1142 AACTGGTTCGGTGGCGTCAATCCGTCCTTACTACCGCGCTTGTATGTCTCCGTCA 1201  
Db 4536 AGCTGGTTCGATTTGGGATCAATAGCCGTGACCTACTACCGCGTCTTGAGTGTCCGTCA 4595  
Qy 1202 TCCGACCACTGGTGTGAGTGTGTCGTGTGGCACTGACGCGCTCATGACCGGCTTACCG 1261  
Db 4596 TCCCGACCACTGGGAGTGTGTCGTGTGGCAACCGATGCGCTCATGACCGGCTATACCG 4655  
Qy 1262 GCGACTTCGATTCGGGTATAGATGCAACACGCTGTCTACCCAGACAGTGCATTCAGCC 1321  
Db 4656 GCGACTTCGATTCGGGTATAGATGCAATACGCTGTCTACCCAGACAGTGCATTCAGCC 4715  
Qy 1322 TTGACCCCTACTTTCACCATTTAGACATCAACGCTTCCCGAGGATGCTGTCTCCCGTACTC 1381  
Db 4716 TTGACCCCTACTTTCACCATTTAGACATCAACGCTTCCCGAGGATGCTGTCTCCCGTACTC 4775  
Qy 1382 AACGTGGGGTAGGACTGGCAGAGGAGACCGAGCATCTACAGATTTGTGGCAACCGGGG 1441  
Db 4776 AACGTGGGGTAGGACTGGCAGAGGAGACCGAGCATCTACAGATTTGTGGCAACCGGGG 4835  
Qy 1442 AGCGTCTCTCTGCGATGTTTGGTCTGCTCTGCGAGTGTATGACGCGGGTGTG 1501  
Db 4836 AGCGTCTCTCTGCGATGTTTGGTCTGCTCTGCGAGTGTATGACGCGGGTGTG 4895  
Qy 1502 CTTGGTATGAGCTTACCGCCCGGAGACCAACAGTGTAGGCTTACGAGCATATCAACACCC 1561  
Db 4896 CTTGGTATGAGCTTACCGCCCGGAGACCAACAGTGTAGGCTTACGAGCATATCAACACCC 4955  
Qy 1562 CGGAGCTTCCGCTGTGCAAGACATCTTGAATTTTGGAGGCGCTTTTACCGGTCTCA 1621  
Db 4956 CGGAGCTTCCGCTGTGCAAGACATCTTGAATTTTGGAGGCGCTTTTACAGGCTCA 5015  
Qy 1622 CCACATAGAGCGCCACTTCTATCCAGACCAAGCAGAGTGGGGAACCTTCCCTATC 1681  
Db 5016 CTATATAGATGGCCACTTCTATCCAGACCAAGCAGAGTGGGGAACCTTCCCTATC 5075  
Qy 1682 TGTAGTGTACCAAGCAGCGCTGTGGCTAGAGCTCAAGCCCTCCCGCTCGTGGGACC 1741  
Db 5076 TGTAGTGTACCAAGCAGCGCTGTGGCTAGAGCTCAAGCCCTCCCGCTCGTGGGACC 5135  
Qy 1742 AGATGTGGAGTGTGTATCGCTCAAGCCCAACCTTCAATGGGCCCAACACTTCTCTAT 1801  
Db 5136 AGATGTGGAGTGTGTATCGCTCAAGCCCAACCTTCAATGGGCCCAACACTTCTCTAT 5195  
Qy 1802 ATAGACTGGCGCTGTCCAGATGAAGTCACTGACGCAACCCAGTCAACCAAGTATATCA 1861  
Db 5196 ACAGACTGGCGCTGTTCAGATGAAGTCACTGACGCAACCCAGTCAACCAATATATCA 5255  
Qy 1862 TGACATGTATGTGGCTGACCTGAGTGTGTACGAGTACCTGGGTGCTGTTGGCGGG 1921  
Db 5256 TGACATGTATGTGGCGGACCTGAGTGTGTACGAGACCTGGGTGCTGTTGGCGGG 5315  
Qy 1922 TTCTGCTGCTTGGCGCGTATTTGCTATFCAAGGCTGCGGTGCTATAGTAGGTAGGA 1981  
Db 5316 TCCTGCTGCTTGGCGCGTATTTGCTGTCACAGGCTGCGGTGCTATAGTAGGTAGGA 5375  
Qy 1982 TTGCTTGTCCGGAAGCGGCAATCATACCGCAGGGAAGTCTCTACCGGAGTTCG 2041  
Db 5376 TCCTGCTTGTCCGGAAGCGGCAATCATACCGCAGGGAAGTCTCTACCGGAGTTCG 5435  
Qy 2042 ATGAAATGGAAGTGTCT 2059  
Db 5436 ATGAGATGGAAGTGTCT 5453

RESULT 14  
AAT12710  
ID AAT12710 standard; cdna; 9401 BP.  
XX  
AC AAT12710;

XX 25-MAR-2003 (revised)  
DT 15-MAY-1996 (first entry)  
XX Hepatitis C virus polyprotein.  
XX Non-A non-B hepatitis virus; NANBHV; HCV; antigen; detection; diagnosis;  
KW antibodies; ds.  
XX Hepatitis C virus.  
XX  
PH Key Location/Qualifiers  
FT CDS 342..9378  
FT /\*tag= a  
XX  
PN BP693687-A1.  
XX  
XX 24-JAN-1996.  
XX  
XX 03-APR-1991; 95BP-00114016.  
XX  
XX 04-APR-1990; 90US-00504352.  
XX  
XX (CHIR ) CHIRON CORP.  
XX  
XX Houghton M, Choo Q, Kuo G;  
PI  
XX WPI; 1996-117956/13.  
DR  
XX P-PSDB; AAR90931.  
XX  
XX Combinations of synthetic Hepatitis C Virus antigens - provide more  
XX effective diagnosis of Non-A, Non-B Hepatitis.  
XX  
XX Disclosure; Fig 1(A-Y); 53pp; English.  
XX  
XX The combination comprises an HCV antigen from the C domain (pref. C22 -  
XX AAR90936) and at least one HCV antigen from the NS3 (pref. C33c -  
XX AAR90932), NS4 (pref. C100 - AAR90933), S (pref. S2 - AAR90935) or NS5  
XX (AAR90934) domain. The antigens may in the form of a fusion protein, a  
XX simple physical mixture, or the individual antigens commonly bound to a  
XX solid matrix. They are pref. prepd. by recombinant DNA techniques  
XX (primers are given in AAT12711-T12716), but can be synthesised or  
XX isolated from HCV using affinity chromatography. (Updated on 25-MAR-2003  
XX to correct PF field.)  
XX  
XX Sequence 9401 BP; 1883 A; 2673 G; 1985 T; 0 U; 0 Other;  
SQ  
Query Match 86.7%; Score 1786; DB 2; Length 9401;  
Best Local Similarity 91.7%; Pred. No. 0;  
Matches 1888; Conservative 0; Mismatches 170; Indels 0; Gaps 0;  
Qy 2 TGGCGCTATACGGCTATGCCAGCAGACAGAGGGGCTTTTGGGATGCATATCACCA 61  
Db 3418 TGGCGCTATACGGCTATGCCAGCAGACAGAGGGGCTTTTGGGATGCATATCACCA 3477  
Qy 62 GCTTGACCGCGCGGACAAAAACAGGTGGAGGTGAGGTTGAGATCGTCAACTGCTG 121  
Db 3478 GCTTACTGCGCGGACAAAAACAGGTGGAGGTGAGGTTGAGATCGTCAACTGCTG 3537  
Qy 122 CCAGACTTTCTTGGCAACCTGCATTAACGGGTCTTGTATCCAGATGTACCAATGTGGACC 181  
Db 3538 CCAGAACTTCTTGGCAACCTGCATTAACGGGTCTTGTATCCAGATGTACCAATGTAGACC 3597  
Qy 182 GAACAAGGACCACTTGGTCCATTAAGGGTCTTGTATCCAGATGTACCAATGTGGACC 241  
Db 3598 GAACAAGGACCACTTGGTCCATTAAGGGTCTTGTATCCAGATGTACCAATGTAGACC 3657  
Qy 242 AAGACCTCGTAGGCTGGCGCGCTCCCAAGGTGCCCGCTCATTAACACCATGCACTTGG 301  
Db 3658 AAGACCTCGTAGGCTGGCGCGCTCCCAAGGTGCCCGCTCATTAACACCATGCACTTGG 3717  
Qy 302 GCTCTCGGACCTTACTGTTGTCACGAGGACGCCGATGTCTTCTGTGCGCGGACGGG 361



3718 GCTCCTCGGACCTTTTACCTGCTACAGAGGACGCGCATGTCTATCCCGTGGCGCGGG 3777  
QY  
362 GTGATGGAGGGGAGCCCTGCTTTGCGCCCGGCGCTATCTCTTACTTTGAAGGCTCTCGG 421  
Db  
3778 GTGATAGCAGGGGAGCCCTGCTGTGCGCCCGGCGCATTTTCTTACTTTGAAGGCTCTCGG 3837  
QY  
422 GAGGCGCTCTGCTGTGCGCCCGGAGGACATGCGTAGGAGATATTACAGAGCGCGGTATGCA 481  
Db  
3838 GGGGTCCGCTGTTGTGCGCCCGGAGGACGCGCTGGGCATATTAGGCGCGCGGTGCA 3897  
QY  
482 CCCGTGAGTGGCTAAGCGGTGGAATTCATCCCGGTAGAGAGCTTTAGAGACAACCATGA 541  
Db  
3898 CCGGTGAGTGGCTAAGCGGTGGAATTTATCCCTGTGAGAACTTAGAGACAACCATGA 3957  
QY  
542 GGTCCCGGTGTTCTCAGACAATCTCTCCCGACAGCAGTGGCCCGAGCTACCAAGTGG 601  
Db  
3958 GGTCCCGGTGTTTACAGGATAATCTCTCTCCACGATAGTGGCCCGAGAGCTTTCCAGGTGG 4017  
QY  
602 CCCACTGTGATGCTCCACCGGAGCGGTAAAGAGCAACCAAGGTCCCGGCGCATACGCGAG 661  
Db  
4018 CTCACCTCCATGCTCCACAGGAGCGGCAAAAGCACCAAGGTCCCGCTGCTATGCGAG 4077  
QY  
662 CTCAGGCTTACAGGTGCTGCTCAACCCCTCCGTTGCTGCAACAATGGCTTTGGTG 721  
Db  
4078 CTCAGGCTTAAAGGTGCTAGTACTCAACCCCTCTGTTGCTGCAACACTGGGCTTTGGTG 4137  
QY  
722 CTTACATGTCGAAGGCCATGGATGATCTTAACATCAGGACTGGGTTGAGGACAAATTA 781  
Db  
4138 CTTACATGTCGAAGGCTCATGGATCGATCTTAACATCAGGACCGGGGTGAGAACAATTA 4197  
QY  
782 CTACTGGCAGCCGATACGTTATTCACCTAGCGCAAGTTCTTGGCCGAGCGGGTGT 841  
Db  
4198 CCACTGGCAGCCCATCACTGACTCCACCTACCGGCAAGTTCTTGGCGAGCGGGTGT 4257  
QY  
842 CAGGGGTGCTTATGACATAATTTGTGACAGGTGCCATCTCCACGAGTGAACAATCA 901  
Db  
4258 CGGGGGCGCTTATGACATAATTTGTGACAGTGCACCTCCACGATGCCACATCA 4317  
QY  
902 TCTTGGGATGACATGCTGCTTACCAAGCAGAGACCGGGGGGAGACTGACTGTGC 961  
Db  
4318 TCTTGGGATGCGCATGCTGCTTGAACAAGCAGAGACTCGGGGGGAGACTGTTGTGC 4377  
QY  
962 TCGCCACCGCTACCCCTCCGGGCTCCGCTCACTGTGCCCATCTTAACATCGAGGAGTTG 1021  
Db  
4378 TCGCCACCGCCACCCCTCCGGGCTCCGCTCACTGTGCCCATCCACATCGAGGAGTTG 4437  
QY  
1022 CTCTGTCCACTACCGGAGATCCCTTTTATGGCAAGGTATTCCTTGTGAAGCAATTA 1081  
Db  
4438 CTCTGTCCACACCGGAGATCCCTTTTACGGCAAGGTATTCCTTGTGAAGTAATCA 4497  
QY  
1082 AGGGGGGAGACATCTCATCTTCTGCCACTCAAGAAGAGTGCAGAGCTCGCGCAA 1141  
Db  
4498 AGGGGGGAGACATCTCATCTTCTGCCACTCAAGAAGAGTGCAGAGCTCGCGCAA 4557  
QY  
1142 AACTGTGTCGGTGGGCGTCAATCGCGGTGCTTACTACCGCGCTTGTATGTCCGTC 1201  
Db  
4558 AGCTGTGTCGATGGGCAATCAATCGCGTCTACTACCGCGCTTGTATGTCCGTC 4617  
QY  
1202 TCCCGACCAAGTGTGAGCTGCTGCTGGCAATCAACCGCTCATGACCGGCTTACCG 1261  
Db  
4618 TCCCGACCAAGTGTGAGCTGCTGCTGGCAATCAACCGCTCATGACCGGCTTACCG 4677  
QY  
1262 GCGACTTCGATTCGGTATAGATGCAACACGTGTGTACCCAGACAGCTCGACTTCAGCC 1321  
Db  
4678 GCGACTTCGATTCGGTATAGATGCAACACGTGTGTACCCAGACAGCTCGACTTCAGCC 4737  
QY  
1322 TTGACCTTACCTTACCACTTGAACAATCACTGTCCCGAGGATGCTGTCTCCGCTACTC 1381  
Db  
4738 TTGACCTTACCTTACCACTTGAACAATCACTGTCCCGAGGATGCTGTCTCCGCTACTC 4797  
QY  
1382 AAGTCCGGGTAGGACTGGCAGAGGAGCCAGGACATCTACAGATTTGTGGCAGCGGGG 1441  
Db  
4798 AAGTCCGGGTAGGACTGGCAGAGGAGCCAGGACATCTACAGATTTGTGGCAGCGGGG 4857

QY 1442 AGCGTCTTCTGGCATGTTTGAATGCTGCTCTCTCGAGTGTATGACGCGGTTGTG 1501  
Db 4858 AGCGCCCTCCGGCATGTTTGAATGCTGCTCTCTCGAGTGTATGACGCGGTTGTG 4917  
QY 1502 CTTGATATGAGCTTACGCGCCCGAGACCAAGTGTAGGCTACGAGCATACATGAACACCC 1561  
Db 4918 CTTGATATGAGCTACGCGCCCGAGACCAAGTGTAGGCTACGAGCATACATGAACACCC 4977  
QY 1562 CGGACTTCCCGTGTGCCAGACCATCTTGAATTTTGGAGGGGCTCTTTACAGGCTTCA 1621  
Db 4978 CGGAGCTTCCCGTGTGCCAGACCATCTTGAATTTTGGAGGGGCTCTTTACAGGCTTCA 5037  
QY 1622 CCCACATAGACGCGCCACTTCTTATCCAGACCAAGTGTAGGCTACGAGCAACCTTCCCTATC 1681  
Db 5038 CTATATATAGTCCCATCTTCTTATCCAGACCAAGTGTAGGCTACGAGCAACCTTCCCTATC 5097  
QY 1682 TGGTAGCGTACCAAGCCACCGTGTGCTGAGCTCAAGCCCTTCCCGCTCGTGGGAGCC 1741  
Db 5098 TGGTAGCGTACCAAGCCACCGTGTGCTGAGCTCAAGCCCTTCCCGCTCGTGGGAGCC 5157  
QY 1742 AGATGTGAACTGCTTGTATGCTCTCAAGCCACCGTGTGAGCTCAAGCCCTTCCCGCTCGTGTAT 1801  
Db 5158 AGATGTGAACTGCTTGTATGCTCTCAAGCCACCGTGTGAGCTCAAGCCCTTCCCGCTCGTGTAT 5217  
QY 1802 ATAGACTGGCGCTGTCCAGAAATGAAGTCAAGCTCAAGCCACCGTGTGAGCTCAAGCTATATCA 1861  
Db 5218 ACAGACTGGCGCTGTTCAGAAATGAATCAAGCTCAAGCCACCGTGTGAGCTCAAGCTATATCA 5277  
QY 1862 TGACATGATGTGGCTGACCTGAGGCTGTGAGCTCAAGCTCAAGCTGTGCTGTGCTGTGCGG 1921  
Db 5278 TGACATGATGTGGCTGACCTGAGGCTGTGAGCTCAAGCTCAAGCTGTGCTGTGCTGTGCGG 5337  
QY 1922 TTCTGGCTGCTTGGCGCGCTATTGCTTATCCAGAGTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 1981  
Db 5338 TCCTGGCTGCTTGGCGCGCTATTGCTTATCCAGAGTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 5397  
QY 1982 TTGCTCTTCTCGGAAAGCGGCAATCATACCCGACAGGAAAGTCTCTTACCGGAGTTGCTG 2041  
Db 5398 TCGTCTTCTCGGAAAGCGGCAATCATACCTGACAGGAAAGTCTCTTACCGGAGTTGCTG 5457  
QY 2042 ATGAATGGAAGTGTCT 2059  
Db 5458 ATGAGATGGAAGTGTCT 5475

## RESULT.15

AAT99981  
ID AAT99981 standard; DNA; 9401 BP.

AC AAT99981;  
XX  
DT 25-MAR-2003 (revised)  
DT 16-MAR-1998 (first entry)  
XX HCV polyprotein coding sequence.

PCR primer; amplify; HCV; hepatitis c virus; antigen combination; NS3;  
C domain; S domain; NS5; HCV polyprotein; anti-HCV antibody; detection;  
NS4; DB.

Hepatitis C virus.

Key Location/Qualifiers  
CDS 342..9377  
/\*tag= a

US5683864-A.

04-NOV-1997.

07-JUL-1992; 92US-00910760.

XX



PR 18-NOV-1987; 87US-00122714.  
 PR 30-DEC-1987; 87US-00139886.  
 PR 26-FEB-1988; 88US-00161072.  
 PR 06-MAY-1988; 88US-00191263.  
 PR 26-OCT-1988; 88US-00263584.  
 PR 14-NOV-1988; 88US-00271450.  
 PR 17-MAR-1989; 89US-00325338.  
 PR 20-APR-1989; 89US-00341334.  
 PR 21-APR-1989; 89US-00353896.  
 PR 18-MAY-1989; 89US-00355002.  
 PR 04-APR-1990; 90US-00504352.  
 XX (CHIR ) CHIRON CORP.  
 XX  
 PI Kuo G, Houghton M, Choo Q;  
 XX  
 DR WPI: 1997-548976/50.  
 DR P-PSDB; AAW34480.  
 XX  
 PT Combination of three hepatitis C virus antigens - used for detection of  
 PT specific antibodies to diagnose infection.  
 XX  
 PS Disclosure; Col 25-46; 57pp; English.  
 XX  
 CC This sequence represents the Hepatitis C virus polypeptide coding  
 CC sequence. Fragments of this sequence can be amplified and used in the  
 CC combination of HCV antigens of the invention. The HCV antigen combination  
 CC comprises an antigen (Ag1) comprising the C domain (i.e. amino acids (aa)  
 CC 1-120 of the HCV polypeptide), or its immunologically reactive fragment  
 CC containing at least 8 aa. It also comprises two additional antigens from  
 CC two different polypeptide domains, including at least 8 aa from the NS3,  
 CC NS4, S or NS5 domains of the polypeptide, corresponding, respectively, to  
 CC aa 1050-1640; 1640-2000; 120-400 and 2000-3011 of the HCV polypeptide.  
 CC Alternatively, Ag1 contains at least 8 aa from the 1-122 or 9-177 aa  
 CC regions of the HCV polypeptide. These antigen combinations are used  
 CC diagnostically to detect anti-HCV antibodies, using any standard  
 CC immunoassay format. These antigen combinations have a broader range of  
 CC reactivity with antibodies than any antigen individually. (Updated on 25-  
 CC MAR-2003 to correct PR field.)  
 XX  
 SQ Sequence 9401 BP; 1883 A; 2860 C; 2673 G; 1985 T; 0 U; 0 Other;  
 Query Match 86.7%; Score 1786; DB 2; Length 9401;  
 Best Local Similarity 91.7%; Pred. No. 0;  
 Matches 1888; Conservative 0; Mismatches 170; Indels 0; Gaps 0;  
 2 TGGCGCTATCAGCGCTATGCCAGCAGACAGAGGGGCTTTGGGATGCATTAATCACCA 61  
 3418 TGGCGGCCATCAGCGCGTAGCGCCAGCAGACAGAGGGGCTTCTAGGGTGATTAATCACCA 3477  
 62 GCTTGACCGCGCGGACAAAACAGGTGAGGGTGAGGTTCAGATCGTGCAACTGCTG 121  
 3478 GCCTAAGTGGCGGGACAAAACAGGTGAGGGTGAGGTTCAGATCGTGCAACTGCTG 3537  
 122 CCAGACTTCTTGGCAACTGCAATTAACGGGGTGTTGGAGATGCTTACCATGAGAGCGG 181  
 3538 CCCAAACCTTCTTGGCAACTGCAATTAACGGGGTGTTGGAGATGCTTACCATGAGAGCGG 3597  
 182 GRACAGGACCATTCGCTGCACTTAAGGTCTCTGTTATCCAGATGTACACCAATGTGAGCC 241  
 3598 GAACGAGGACCATTCGCTGCACTTAAGGTCTCTGTTATCCAGATGTACACCAATGTGAGCC 3657  
 242 AAGACCTCGTAGGCTGCGCGGCTCCCAAGGTGCGGCTCATTAACACCATGACATTTGGC 301  
 3658 AAGACCTTGTGGCTGCGCGGCTCCGCAAGGTAGCGCTCATTAACACCATGACATTTGGC 3717  
 302 GCTCTCTGGACCTTTACCTGGTCAAGGACACCGCATGTCATTCCTGTGCGCGGACGGG 361  
 3718 GCTCTCTGGACCTTTACCTGGTCAAGGACACCGCATGTCATTCCTGTGCGCGGACGGG 3777  
 362 GTGATGCGAGGCGGACGCTGCTTTGCGCGGCGCTATCTTACTTTGAAGGCTCTCTGG 421  
 3778 GTGATGCGAGGCGGACGCTGCTGCTGCGCGGCGCTATCTTACTTTGAAGGCTCTCTGG 3837

QY 422 GAGGCCCTCTGTGTGTCGCCCGCAGGACATGCCGTAGGCATATTCAGAGCGCGGTATGCA 481  
 DB 3838 GGGGTCCGTGTGTGTGTCGCCCGCAGGACACCGGTGGGCATATTTAGGGCGCGGTGTGCA 3897  
 QY 482 CCGGTGGATGCTAAGCGGTGGACTTCATCCCGTAGAGACTTAGAGACACCATGA 541  
 DB 3898 CCGGTGGATGCTAAGCGGTGGACTTCATCCCGTAGAGACTTAGAGACACCATGA 3957  
 QY 542 GGTCCCCGGTGTCTCAGACAACTCTCCACCACAGCAGTGCACAGAGCTACCAAGTGG 601  
 DB 3958 GGTCCCCGGTGTCTCAGGATACTCTCTCCACAGTGTGCCCCAGAGCTTCCAGGTGG 4017  
 QY 602 CCCACTGTCATGCTCCACCGCAGCGGTAAGAGCAACCAAGGTCCCGCGCATACGCAG 661  
 DB 4018 CTCACCTCATGCTCCACAGCAGCGGCAAAAGCACCAGGTCCCGCTGCATATGCAG 4077  
 QY 662 CTCAGGCTACAGGTGCTGCTCAACCCCTCTGCTGCTGCTGCAACATGGGCTTTGGTG 721  
 DB 4078 CTCAGGCTATAGGTGCTGCTCAACCCCTCTGCTGCTGCTGCAACATGGGCTTTGGTG 4137  
 QY 722 CTTACATGTCCAAGGCCCATGGATGATCTTAACATCAGGACTGGGGTGAGACAAATTA 781  
 DB 4138 CTTACATGTCCAAGGCTCATGGATGATCTTAACATCAGGACTGGGGTGAGACAAATTA 4197  
 QY 782 CTACTGGCAGCCGATCAGTATTCACCTACCGCAAGTTCCTTGCAGCGCGGTGT 841  
 DB 4198 CCACTGGCAGCCCATCAGTACTCCACCTACCGCAAGTTCCTTGCAGCGCGGTGT 4257  
 QY 842 CAGGGGTGCTTATGACATAAATTTGTGACAGTGCACCTCCAGGATGCAACATCCA 901  
 DB 4258 CGGGGGCGCTTATGACATAAATTTGTGACAGTGCACCTCCAGGATGCAACATCCA 4317  
 QY 902 TCTTGGGCTTGGCAGTCTGCTTGCACAGAGAGACCGCGGGGGAGACTGACATGTGC 961  
 DB 4318 TCTTGGGCTTGGCAGTCTGCTTGCACAGAGAGACTGCGGGGGAGACTGACATGTGC 4377  
 QY 962 TCGCCACCGCTACCCCTCGGGCTCGGTCACTGTGCCCATCTTAACATCGAGGAGTTG 1021  
 DB 4378 TCGCCACCGCCACCCCTCGGGCTCGGTCACTGTGCCCATCTTAACATCGAGGAGTTG 4437  
 QY 1022 CTCTGTCACTACCGAGAGATCCCTTTTATGGCAGGCTATTCCTTGAAGCAATTA 1081  
 DB 4438 CTCTGTCACTACCGAGAGATCCCTTTTATGGCAGGCTATTCCTTGAAGCAATTA 4497  
 QY 1082 AGGGGGGAGACATCTCATCTTCTGCACTCAAAAGAAAGTGCAGAGCTCGCGCAA 1141  
 DB 4498 AGGGGGGAGACATCTCATCTTCTGCACTCAAAAGAAAGTGCAGAGCTCGCGCAA 4557  
 QY 1142 AACTGTGCGGTGGGCGTCAATGCGGTGGCTTACTACCGCGGCTTGTATGTGTCGTCA 1201  
 DB 4558 AGTGTGCGCATTTGGGCATCAATGCGGTGGCTTACTACCGCGGCTTGTATGTGTCGTCA 4617  
 QY 1202 TCCGACAGTGTGAGCTGTGCTGCGCAACTGACCGCTCATGACGGGTTTACCG 1261  
 DB 4618 TCCGACAGTGTGAGCTGTGCTGCGCAACTGACCGCTCATGACGGGTTTACCG 4677  
 QY 1262 GCGACTTCGATTCGGGTAGACTGCAACACGTGTGTCAACCGAGAGCTGACTTCAGGC 1321  
 DB 4678 GCGACTTCGATTCGGGTAGACTGCAACACGTGTGTCAACCGAGAGCTGACTTCAGGC 4737  
 QY 1322 TTGACCTTACCTTACCATTTAGACAAATCACGCTTCCCGAGGATGCTGTCTCCGTA 1381  
 DB 4738 TTGACCTTACCTTACCATTTAGACAAATCACGCTTCCCGAGGATGCTGTCTCCGTA 4797  
 QY 1382 AACGTGGGTTAGGACTGGCAGAGGAGCAAGGATCTACAGATTTGTGGACCCCGGGG 1441  
 DB 4798 AACGTGGGTTAGGACTGGCAGAGGAGCAAGGATCTACAGATTTGTGGACCCCGGGG 4857  
 QY 1442 AGCGTCTTCTGGCATGTTTGTGCTGCTGCGAGTGTATGATGCGGGGTTG 1501  
 DB 4858 AGCGTCTTCTGGCATGTTTGTGCTGCTGCGAGTGTATGATGCGGGGTTG 4917

Mon Feb 28 11:40:44 2005

QY	1502	CTTGGTATGAGCTTAGCCCGCCGAGACACAGTTAGGCTACGACATACATGAACACC	1561
Db	4918	CTTGGTATGAGCTACGCGCCGCGAGACTACAGTTAGGCTACGAGCTACATGAACACC	4977
QY	1562	CGGGACTTCCCGTGTGCCAAGACCATCTTGAATTTGGAGGGCGCTTTACGGGTCTCA	1621
Db	4978	CGGGGCTTCCCGTGTGCCAGGACCATCTTGAATTTGGAGGGCGCTTTACAGCCCTCA	5037
QY	1622	CCACATAGACGCCACTTCTCTATCCAGACAAAGCAGAGTGGGGAACCTTCCCTATC	1681
Db	5038	CTCATATAGATGCCCACTTCTCTATCCAGACAAAGCAGAGTGGGGAACCTTCCCTATC	5097
QY	1682	TGCTAGCGTACCAAGCCACCGTGTGCGCTAGAGCTCAAGCCCTTCCCGTGTGGGACC	1741
Db	5098	TGCTAGCGTACCAAGCCACCGTGTGCGCTAGAGCTCAAGCCCTTCCCGTGTGGGACC	5157
QY	1742	AGATGTGGAAGTCTTGATCCGCTCAAGCCACCCCTCCATGGGCCACACCTCTGCTAT	1801
Db	5158	AGATGTGGAAGTCTTGATCCGCTCAAGCCACCCCTCCATGGGCCACACCCCTGCTAT	5217
QY	1802	ATAGACTGGGCGCTGTCCAGAATGAAGTCAACCTGACGCAACCCAGTCAACAAATATCA	1861
Db	5218	ACAGACTGGGCGCTGTTCAGATGAATCAACCTGACGCAACCCAGTCAACAAATATCA	5277
QY	1862	TGACATGTATGTGCGGTGACCTGGAGGTGCTACGAGTACCTGGGTGCTCGTTGGGGCG	1921
Db	5278	TGACATGTATGTGCGGTGACCTGGAGGTGCTACGAGTACCTGGGTGCTCGTTGGGGCG	5337
QY	1922	TTCTGGCTGCTTTGGCGCGTATTGCTATCCACAGGCTGCGTGGTCTAGTAGGTAGGA	1981
Db	5338	TCCTGGCTGCTTTGGCGCGTATTGCTATCCACAGGCTGCGTGGTCTAGTAGGTAGGA	5397
QY	1982	TTGTCTTTGTCGGGAAGCCGGCAATCATACCGACAGGGAAGTCTCTACCGGGAGTTCCG	2041
Db	5398	TCGTCTTTGTCGGGAAGCCGGCAATCATACCGACAGGGAAGTCTCTACCGGGAGTTCCG	5457
QY	2042	ATGAATGGAAGAGTGCT	2059
Db	5458	ATGAGATGGAAGAGTGCT	5475

Search completed: February 27, 2005, 18:30:06  
Job time : 1064 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 27, 2005, 17:06:13 ; Search time 370 Seconds  
(without alignments)  
9114.504 Million cell updates/sec

Title: US-09-930-591-1  
Perfect score: 2061  
Sequence: 1 atggcgcttatcacggccta.....atgaatggaagagtgtga 2061

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents NA: \*  
1: /cgn2\_6/prodata/1/ina/5A COMB.seq: \*  
2: /cgn2\_6/prodata/1/ina/5B COMB.seq: \*  
3: /cgn2\_6/prodata/1/ina/6A COMB.seq: \*  
4: /cgn2\_6/prodata/1/ina/6B COMB.seq: \*  
5: /cgn2\_6/prodata/1/ina/PCTUS COMB.seq: \*  
6: /cgn2\_6/prodata/1/ina/backfiles1.seq: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1786	86.7	7310	3	US-08-444-818-74
2	1786	86.7	9379	3	US-09-388-874-1
3	1786	86.7	9379	4	US-09-916-359-1
4	1786	86.7	9401	1	US-07-910-760-9
5	1786	86.7	9401	1	US-08-440-519-9
6	1786	86.7	9401	3	US-08-440-549-9
7	1786	86.7	9401	3	US-08-823-895A-25
8	1785.2	86.6	9401	5	PCT-US91-02225-9
9	1784.4	86.6	6785	3	US-08-444-818-65
10	1784.4	86.6	8316	3	US-08-444-818-88
11	1784.4	86.6	8987	3	US-08-444-818-137
12	1784.4	86.6	9185	3	US-08-444-818-122
13	1784.4	86.6	9379	3	US-08-444-818-123
14	1784.4	86.6	9379	3	US-08-444-818-176
15	1781.2	86.4	2058	4	US-09-881-239-2
16	1781.2	86.4	5360	3	US-08-444-818-53
17	1779.6	86.3	2058	4	US-09-881-654-1
18	1779.6	86.3	2058	4	US-10-637-323-1
19	1774.8	86.1	12980	3	US-08-811-566-5
20	1774.8	86.1	12980	3	US-09-034-756-5
21	1773.2	86.0	9646	3	US-08-811-566-1
22	1773.2	86.0	9646	3	US-09-034-756-1
23	1765.2	85.6	9599	3	US-09-014-416-2
24	1765.2	85.6	9599	3	US-09-014-416-6
25	1762	85.5	9416	3	US-08-823-895A-26
26	1762	85.5	9416	4	US-10-104-966-13
27	1750.8	84.9	9401	2	US-08-432-693-1

28 1750.8 84.9 9416 3 US-08-811-566-19 Sequence 19, Appl  
29 1750.8 84.9 9416 3 US-09-034-756-19 Sequence 19, Appl  
30 1749.2 84.9 9365 4 US-09-827-688-7 Sequence 7, Appl  
31 1557.8 75.6 2064 1 US-08-350-884-69 Sequence 69, Appl  
32 1557.8 75.6 2064 1 US-08-440-548-69 Sequence 69, Appl  
33 1557.8 75.6 2064 1 US-08-709-173-69 Sequence 69, Appl  
34 1557.8 75.6 2064 2 US-08-709-177-69 Sequence 69, Appl  
35 1557.8 75.6 2523 1 US-08-350-884-85 Sequence 85, Appl  
36 1557.8 75.6 2523 1 US-08-440-548-85 Sequence 85, Appl  
37 1557.8 75.6 2523 1 US-08-709-173-85 Sequence 85, Appl  
38 1557.8 75.6 2523 2 US-08-709-177-85 Sequence 85, Appl  
39 1546 75.0 1932 3 US-09-128-314-1 Sequence 1, Appl  
40 1546 75.0 8157 3 US-09-128-314-3 Sequence 3, Appl  
41 1398.4 67.9 9413 4 US-09-827-688-6 Sequence 6, Appl  
42 1381.2 67.0 7989 4 US-09-539-601-10 Sequence 10, Appl  
43 1381.2 67.0 8001 4 US-09-539-601-7 Sequence 7, Appl  
44 1379.6 66.9 8001 4 US-09-539-601-22 Sequence 22, Appl  
45 1379.2 66.9 8638 4 US-10-029-907-7 Sequence 7, Appl

ALIGNMENTS

RESULT 1  
US-08-444-818-74  
; Sequence 74, Application US/08444818  
; Patent No. 6150087  
; GENERAL INFORMATION:  
; APPLICANT: Chien, David Y.  
; APPLICANT: Rutter, William J.  
; TITLE OF INVENTION: NANBV Diagnostics and Vaccines  
; NUMBER OF SEQUENCES: 777  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Chiron Corporation  
; STREET: 4560 Horton Street  
; CITY: Emeryville  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 94608-2916  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/444,818  
; FILING DATE:  
; CLASSIFICATION: 424  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/08/403,590  
; FILING DATE: 14-MAR-1995  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Harbin, Alisa A.  
; REGISTRATION NUMBER: 33,895  
; REFERENCE/DOCKET NUMBER: 0110.002  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (508)359-3876  
; TELEFAX: (508)359-3885  
; INFORMATION FOR SEQ ID NO: 74:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 7310 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: cDNA  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: 3..7310  
; US-08-444-818-74

Query Match 86.7%; Score 1786; DB 3; Length 7310;  
Best Local Similarity 91.7%; Pred. No. 0;  
Matches 1898; Conservative 0; Mismatches 170; Indels 0; Gaps 0;

QY 2 TGGCGCTATCAGCGCTATGCCAGCAGACAAAGGGCCCTTTGGATGTCATAATCACCA 61  
DB 1729 TGGCGCCATCAGCGGTAAGCCAGCAGACAAAGGGCCCTCTAGGTGTCATAATCACCA 1788  
QY 62 GCTTACCGCGCGGACAAAACCAAGGTGAGGTGAGGTTCAGATCGTGTCAACTGCTG 121  
DB 1789 GCTTAACGTGGCGGACAAAACCAAGGTGAGGTGAGGTTCAGATCGTGTCAACTGCTG 1848  
QY 122 CCCAGACTTCTTGGCAACCTGCATTAACGGGGTGTGTGAGTGTCTACCATGAGAGCG 181  
DB 1849 CCCAAACCTTCTTGGCAACCTGCATTAACGGGGTGTGTGAGTGTCTACCATGAGAGCG 1908  
QY 182 GAACAAGGACCAATGCGTCACTAAGGTCTGTGTTATCCAGATGTACACCAATGTGGAC 241  
DB 1909 GAACGAGGACCAATGCGTCACTAAGGTCTGTGTTATCCAGATGTATACCAATGTAGAC 1968  
QY 242 AAGACCTGTAGGCTGGCGGCTGCCCAAGGTGCCCGCTCAATTAACACCATGCACTTGG 301  
DB 1969 AAGACCTGTGGCTGGCGGCTGCCCAAGGTAGCGGCTCAATTAACACCATGCACTTGG 2028  
QY 302 GCTCTCGACCTTACCTGTGTACAGGACGCGGATGTCTTCTTAAAGGCTCCTCGG 361  
DB 2029 GCTCTCGACCTTACCTGTGTACAGGACGCGGATGTCTTCTTAAAGGCTCCTCGG 2088  
QY 362 GTGATGGCAGGGGACGCTGCTTGGCCCGGCTTCTTCTTAAAGGCTCCTCGG 421  
DB 2089 GTGATAGAGGGGACGCTGCTTGGCCCGGCTTCTTCTTAAAGGCTCCTCGG 2148  
QY 422 GAGGCGCTGTGTGTCGCCCGGACGACATGCCGTAGGCATATTCAGAGCGCGGTATGCA 481  
DB 2149 GGGTCCGCTGTGTGTCGCCCGGACGACATGCCGTAGGCATATTCAGAGCGCGGTATGCA 2208  
QY 482 CCCGTGAGTGTGTAAGGCGGTGAGTTCATCCCGTAGAGAGCTTACAGACAACTATGA 541  
DB 2209 CCCGTGAGTGTGTAAGGCGGTGAGTTCATCCCGTAGAGAGCTTACAGACAACTATGA 2268  
QY 542 GGTCCCGGTGTCTCAGACAACTCTCCCGACAGCAGTCCCGACAGCTACCAAGTGG 601  
DB 2269 GGTCCCGGTGTCTCAGACAACTCTCCCGACAGCAGTCCCGACAGCTACCAAGTGG 2328  
QY 602 CCCACTGTGATGCTCCACCGGACGCGTAAGAGCACCAAGTCCCGCGCGGATACGAC 661  
DB 2329 CTCACCTCATGCTCCACAGCAGCGGCAAAAGCACCAAGTCCCGCGGTCATATGAC 2388  
QY 662 CTCAGGGCTACAGGTGTGCTCAACCCCTCGTGTCTGCTGCAACAATGGCTTTGGT 721  
DB 2389 CTCAGGGCTATAAGGTGTGCTACTCAACCCCTCTGTGCTGCAACAATGGCTTTGGT 2448  
QY 722 CTTTACATGTCCAAGGCCATGGATGATCTTAACATCAGGACTGGGGTGAGGACAATTA 781  
DB 2449 CTTTACATGTCCAAGGCCATGGATGATCTTAACATCAGGACTGGGGTGAGGACAATTA 2508  
QY 782 CTATGGCAGCCCGATCACTATTCACCTACCGCAAGTTCCTTCCGACGCGGGTGT 841  
DB 2509 CCACGTGGCAGCCCGATCACTATTCACCTACCGCAAGTTCCTTCCGACGCGGGTGT 2568  
QY 842 CAGGGGTCTTATGACATAATTTGTGAGAGTGGCCACTCCACGGATGCAATCA 901  
DB 2569 CAGGGGTCTTATGACATAATTTGTGAGAGTGGCCACTCCACGGATGCAATCA 2628  
QY 902 TCTTGGGCAATGGCACTGTCTTACCAAGCAGAGACCGGGGGGAGACTGACTGTGC 961  
DB 2629 TCTTGGGCAATGGCACTGTCTTACCAAGCAGAGACTGGGGGGGAGACTGACTGTGC 2688  
QY 962 TCGCCACCGCTACCCCTCGGGCTCGGTCACTGTGCCCATCTTAACATCGAGGAGTTG 1021  
DB 2689 TCGCCACCGCTACCCCTCGGGCTCGGTCACTGTGCCCATCTTAACATCGAGGAGTTG 2748  
QY 1022 CTTCTGTCCACTACCGAGAGATCCCTTTTATGGCAAGGCTATCCCTTGAAGCAATTA 1081  
DB 2749 CTTCTGTCCACTACCGAGAGATCCCTTTTATGGCAAGGCTATCCCTTGAAGCAATTA 2808

QY 1082 AGGGGGGAGACATCTCATCTTCTGCCACTCAAGAGAGAGTGGACAGCTCGCGCAA 1141  
DB 2809 AGGGGGGAGACATCTCATCTTCTGCCACTCAAGAGAGAGTGGACAGCTCGCGCAA 2868  
QY 1142 AACTGGTGGCTTGGGCGCTCAATGCGGTGCTTACTACGGCGCTTCAATGTGTCCGTCA 1201  
DB 2869 AGCTGGTGGCTTGGGCGCTCAATGCGGTGCTTACTACGGCGCTTCAATGTGTCCGTCA 2928  
QY 1202 TCCGACAGGTGGTGAAGTGTGCTGCTGGCAACTGACGCGCTCATGACGCGCTTACCG 1261  
DB 2929 TCCGACAGGTGGGCGCTTGTGCTGCTGGCAACCGATGCCCTCATGACGCGCTTACCG 2988  
QY 1262 GCGACTTCGATTCGCTGATAGACTGCAACAGTGTGTCCACGACAGTCTGACTTCAGCC 1321  
DB 2989 GCGACTTCGCTGATAGACTGCAATCGTGTGTCCACGACAGTCTGATTTTCAGCC 3048  
QY 1322 TTGACCTTACCTTACCATTTGAGACATCAAGTGTGCCAGGATGCTGTCTCCGTACTC 1381  
DB 3049 TTGACCTTACCTTACCATTTGAGACATCAAGTGTGCCAGGATGCTGTCTCCGTACTC 3108  
QY 1382 AACGTGGGGTAGGACTGGCAGAGGGAAGCAGGCTCTACAGATTTGTGGCAGCGGG 1441  
DB 3109 AACGTGGGGTAGGACTGGCAGAGGGAAGCAGGCTCTACAGATTTGTGGCAGCGGG 3168  
QY 1442 AGCGTCTTCTGGCATGTTTGAATGCTGTCTCTGCGAGTGTATGACGCGGTTGT 1501  
DB 3169 AGCGTCTTCTGGCATGTTTGAATGCTGTCTCTGCGAGTGTATGACGCGGTTGT 3228  
QY 1502 CTTGTATGAGTTACGCGCGGAGACCAAGTGTAGGTAGGCTAGGACATACATGAACGCC 1561  
DB 3229 CTTGTATGAGTTACGCGCGGAGACCAAGTGTAGGTAGGCTAGGACATACATGAACGCC 3288  
QY 1562 CGGACTTCCCGTGTGCCAAGACCATCTTGAATTTTGGAGGGCTCTTTACGGGTCTCA 1621  
DB 3289 CGGACTTCCCGTGTGCCAAGACCATCTTGAATTTTGGAGGGCTCTTTACGGGTCTCA 3348  
QY 1622 CCCACATAGACGCCCATCTTCTATCCACAGACAAAGCAGAGTGGGAAACCTTCCCTATC 1681  
DB 3349 CTCATATAGATGCCCATCTTCTATCCACAGACAAAGCAGAGTGGGAGAACCTTCTTACC 3408  
QY 1682 TGGTAGGCTACAGGCCACCGTGTGGCTAGAGCTCAAGCCCTCCCGGCTGGGAGC 1741  
DB 3409 TGGTAGGCTACAGGCCACCGTGTGGCTAGGCTCAAGCCCTCCCGGCTGGGAGC 3468  
QY 1742 AGATGTGAAGTGTGTGATCGTCTCAAGCCACCTCTCATGGGCAACACCTCTGTCTAT 1801  
DB 3469 AGATGTGAAGTGTGTGATCGTCTCAAGCCACCTCTCATGGGCAACACCTCTGTCTAT 3528  
QY 1802 ATAGACTGGGCGCTGTCCAGAAATGAAGTCACTGACGACCCAGTCAACCAAGTATCA 1861  
DB 3529 ACAGACTGGGCGCTGTTCAGAAATGAAGTCACTGACGACCCAGTCAACCAAGTATCA 3588  
QY 1862 TGACATGTATGTGGCTGACCTGGAGGTCTCAAGAGTACCTGGGTGCTCGTGGGGGG 1921  
DB 3589 TGACATGTATGTGGCTGACCTGGAGGTCTCAAGAGTACCTGGGTGCTCGTGGGGGG 3648  
QY 1922 TTTCTGGCTGTGTGGCGCGTATTCCTATCCACAGGCTGGGTGCTATAGTAGGTAGGA 1981  
DB 3649 TTTCTGGCTGTGTGGCGCGTATTCCTATCCACAGGCTGGGTGCTATAGTAGGTAGGA 3708  
QY 1982 TTGTCTTGTGGGAAAGCGGCAATCATACCGACAGGGAAGTCTCTTACCGGGAGTTG 2041  
DB 3709 TTGTCTTGTGGGAAAGCGGCAATCATACCGACAGGGAAGTCTCTTACCGGGAGTTG 3768  
QY 2042 ATGAATGGAAGAGTGTCT 2059  
DB 3769 ATGAGATGGAAGAGTGTCT 3786

RESULT 2  
US-09-388-874-1  
; Sequence 1, Application US/09388874  
; Patent No. 6284249

## ; GENERAL INFORMATION:

; APPLICANT: Veronique Barban  
; TITLE OF INVENTION: VACCINE COMPOSITION FOR PREVENTING OR  
; TITLE OF INVENTION: TRGATING C HEPATITIS  
; FILE REFERENCE: PMCF97-03A  
; CURRENT APPLICATION NUMBER: US/09/388,874  
; CURRENT FILING DATE: 1999-09-02  
; EARLIER APPLICATION NUMBER: PCT/FR98/00448  
; EARLIER FILING DATE: 1998-03-06  
; EARLIER APPLICATION NUMBER: 97/02,887  
; EARLIER FILING DATE: 1997-03-06  
; NUMBER OF SEQ ID NOS: 2  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 1  
; LENGTH: 9379  
; TYPE: DNA  
; ORGANISM: Virus  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (320)... (9352)  
US-09-388-874-1

Query Match 86.7%; Score 1786; DB 3; Length 9379;

Best Local Similarity 91.7%; Pred. No. 0; Mismatches 170; Indels 0; Gaps 0;  
Matches 1888; Conservative 0;

QY	2	TGCGCGCTATCACGGCCTATGCCAGCAGACAAAGGGCCCTTTGGGATGCATAATCACCA	61
DB	3396	TGGGGCCATACAGGCGTAGCCCGCAGACAAAGGGCCCTCTAGGGTGCATAATCACCA	3455
QY	62	GCTTGACCGCGCGGGAAGAAACCAAGGTGGAGGGTGAAGTTCAGATCGTGTCAACTGCTG	121
DB	3456	GCTTAACCTGGCGGGACAAACCAAGTGGAGGGTGAAGTTCAGATCGTGTCAACTGCTG	3515
QY	122	CCGAGACTTCTTGGCAACCTGCATTAACGGGGTGTGTGGACTGTCTACCATGGAGCG	181
DB	3516	CCCAACCTTCTTGGCAACCTGCATTAACGGGGTGTGTGGACTGTCTACCATGGAGCG	3575
QY	182	GAACAGGACCATTTGGTCACTTCAAGGGTCTGTATCCAGATGTACCAATGTGGACC	241
DB	3576	GAACAGGACCATTTGGTCACTTCAAGGGTCTGTATCCAGATGTATACCAATGTAGACC	3635
QY	242	AAGACCTCGTAGCTGGCCCGCTCCCAAGGTGCCCGCTCATTAACACCATGCATTTGG	301
DB	3636	AAGACCTTGTGGCTGGCCCGCTCCCAAGGTAGCGCTCATTTGACACCTGCATTTGG	3695
QY	302	GCTCTCGAGACCTTACCTGCTGTCAGAGGACGCCGATGTCACTCTGTGGCGGACGG	361
DB	3696	GCTCTCGAGACCTTACCTGCTGTCAGAGGACGCCGATGTCACTCTGTGGCGGACGG	3755
QY	362	GTGATGGCAGGGCAGCCTGCTTTGCCCGCGCTATCTTACTTTGAAAGGCTCCTCGG	421
DB	3756	GTGATAGAGGGCAGCCTGCTGCTGCCCGCGCCCACTTCTACTTTGAAAGGCTCCTCG	3815
QY	422	GAGGCCCTCTGCTGTGCCCGCAGGACATGCCGTAGGCATATTTAGAGCGCGGTATGCA	481
DB	3816	GGGGTCCGCTGTGTGCGCGCGGGGACGCCGTGGGCATATTTAGGGCGCGGTGCA	3875
QY	482	CCGCTGAGTGGCTAAGGGGTGGACTTCAATCCCGTAGAGCTTAGAGACAACATGA	541
DB	3876	CCGCTGAGTGGCTAAGGGGTGGACTTATATCCCTGTGGAGAACCTTAGAGACAACATGA	3935
QY	542	GGTCCCGGTGTCTCAGACAACCTCTCCCAACAGCAGTCCCGCAGAGCTACCAAGTGG	601
DB	3936	GGTCCCGGTGTCTCAGACAACCTCTCTCCACAGTAGTGCCCGCAGAGCTTCCAGGTGG	3995
QY	602	CCCACTGATGCTCCCAACGGCAGCGGTAAAGACCAACAGGTCCCGCGCGCATACGAC	661
DB	3996	CTCACTTCCATGCTCCCAACAGCAGCGGTAAAGACCAACAGGTCCCGCGGTGCATGCA	4055
QY	662	CTCAGGGCTACAGGTGCTGGTGTCAACCCCTCGTGTGCTGCAACAATGGGCTTTGGT	721
DB	4056	CTCAGGGCTATAAGGTGCTAGTACTCAACCCCTCTGTGTGCTGCAACAATGGGCTTTGGT	4115

QY	722	CTTACATGTCCAAGGCCCATGGATGTGATCTTAAACATCAGGACTGGGGTGGAGACAATTA	781
DB	4116	CTTACATGTCCAAGGCTCATGGATCGATCTTAAACATCAGGACCGGGGTGAGAACATTA	4175
QY	782	CTACTGGCAGCCCGATCAGCTATTCCACTAGCGCAAGTTCCTTGGCGAGCGGGGTGT	841
DB	4176	CCACTGGCAGCCCGATCAGCTATTCCACTAGCGCAAGTTCCTTGGCGAGCGGGGTGT	4235
QY	842	CAGGGGTGTCTTATGACATATAATTTGTGACAGTGCACCTCCACGGATGCAACATCCA	901
DB	4236	CGGGGGCGCTTATGACATATAATTTGTGACAGTGCACCTCCACGGATGCAACATCCA	4295
QY	902	TCTTGGCAATTGGCACTGTCTTTGACCAAGCAGACCGCGGGGGGAGACTGACTGTGC	961
DB	4296	TCTTGGCATCGGCATGTCTTTGACCAAGCAGAGACTGCGGGGGGAGACTGGTGTGC	4355
QY	962	TGCGCACCGCTACCCCTCGGGCTCGTCACTGTGCCCATCTTAACATCGAGGAGTTG	1021
DB	4356	TCGCCACCGCCACCCCTCGGGCTCGTCACTGTGCCCATCTTAACATCGAGGAGTTG	4415
QY	1022	CTCTGTCACTACCGGAGATCCCTTTTATGGCAAGCTATTCCTCTTGAAGCAATTA	1081
DB	4416	CTCTGTCACTACCGGAGATCCCTTTTATGGCAAGCTATTCCTCTTGAAGCAATTA	4475
QY	1082	AGGGGGGAGACATCTCTTCTGCACTCAAGAAGAGTGCAGAGCTGCGCGCAA	1141
DB	4476	AGGGGGGAGACATCTCTTCTGCACTCAAGAAGAGTGCAGAGCTGCGCGCAA	4535
QY	1142	AACTGGTGGCTGGCGCTCAATGGCGCTTACTACCGCGGCTTGTGTGCTGCTCA	1201
DB	4536	AGCTGGTGGCTGGCGCTCAATGGCGCTTACTACCGCGGCTTGTGTGCTGCTCA	4595
QY	1202	TCCCGACAGTGTGCTGCTGCGCAACTGACGCCCTCATGACCGGCTTTACCG	1261
DB	4596	TCCCGACAGTGTGCTGCTGCGCAACTGACGCCCTCATGACCGGCTTTACCG	4655
QY	1262	GGGACTTCTGCTGGCTGATGCACTGCAACACGCTGTGTCAACAGAGCTGCTTACGCC	1321
DB	4656	GGGACTTCTGCTGGCTGATGCACTGCAACACGCTGTGTCAACAGAGCTGCTTACGCC	4715
QY	1322	TTGACCTTACCTTCACTAGAGCAATCAACGCTTCCAGAGTGTGTCTCCGCTATC	1381
DB	4716	TTGACCTTACCTTCACTAGAGCAATCAACGCTTCCAGAGTGTGTCTCCGCTATC	4775
QY	1382	AACTGGGGGTAGGACTGGCAGAGGAGCAGGAGTCTACAGATTTGGGACCGGGG	1441
DB	4776	AACTGGGGGTAGGACTGGCAGAGGAGCAGGAGTCTACAGATTTGGGACCGGGG	4835
QY	1442	AGCGTCTTCTGGCATGTTTGAATCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTG	1501
DB	4836	AGCGTCTTCTGGCATGTTTGAATCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTG	4895
QY	1502	CTTGTGTAGCTTACGCGCGCGGAGACCAAGTGTAGGCTACAGAGTATGAGAACACC	1561
DB	4896	CTTGTGTAGCTTACGCGCGCGGAGACCAAGTGTAGGCTACAGAGTATGAGAACACC	4955
QY	1562	CGGGACTTCCCGGTGCGCAAGACCTTGAATTTGGAGGGCGTCTTTAGGGCTCA	1621
DB	4956	CGGGACTTCCCGGTGCGCAAGACCTTGAATTTGGAGGGCGTCTTTAGGGCTCA	5015
QY	1622	CCCATATAGACGCCCACTTCTTATCCAGACAAAGCAGAGTGGGGAACCTTCCCTATC	1681
DB	5016	CTCATATAGATGCCCACTTCTTATCCAGACAAAGCAGAGTGGGGAACCTTCCCTATC	5075
QY	1682	TGGTAGCTACCAAGCCACCGCTGTGCTGTAGAGCTCAAGCCCTCCCGCTGGGAGC	1741
DB	5076	TGGTAGCTACCAAGCCACCGCTGTGCTGTAGAGCTCAAGCCCTCCCGCTGGGAGC	5135
QY	1742	AGATGTGGAAGTGTGTGCTGTCAAGCCCACTTCAAGCCCACTTCAAGCCCACTTCA	1801
DB	5136	AGATGTGGAAGTGTGTGCTGTCAAGCCCACTTCAAGCCCACTTCAAGCCCACTTCA	5195

QY 1802 ATAGACTGGCGCTGTCAGAAATAGTCAACCTGACGACCCAGTCAACAAAGTATATCA 1861  
DB 5196 ACAGACTGGCGCTGTCAGAAATAGTCAACCTGACGACCCAGTCAACAAATATATCA 5255  
QY 1862 TGACATGATGTCGGCTGACCTGAGGTGCTACGAGTACCTGGGTGCTGTTGGCGGG 1921  
DB 5256 TGACATGATGTCGGCGGACCTGGAGGTGCTACGAGCACTGGGTGCTGTTGGCGGG 5315  
QY 1922 TTCTGCTCTTTGGCGCGGTATGCTATCCACAGGTGCTGCTCATAGTAGTAGGA 1981  
DB 5316 TCTGCTCTTTGGCGCGGTATGCTATCCACAGGTGCTGCTCATAGTAGTAGGA 5375  
QY 1982 TTGCTGTCGGAAGCGGCAATCATACCGGACAGGAAAGTCTCTACCGGAGTTGG 2041  
DB 5376 TCGTCTGTCGGAAGCGGCAATCATACCTGACAGGAAAGTCTCTACCGAGATTGG 5435  
QY 2042 ATGAATGAAGAGTGCT 2059  
DB 5436 ATGAGATGAAGAGTGCT 5453

RESULT 3  
US-09-916-359-1  
; Sequence 1, Application US/09916359  
; Patent No. 6538123  
; GENERAL INFORMATION:  
; APPLICANT: Veronique Barban  
; TITLE OF INVENTION: VACCINE COMPOSITION FOR PREVENTING OR  
; TITLE OF INVENTION: TREATING C HEPATITIS  
; FILE REFERENCE: PMPF97-03A  
; CURRENT APPLICATION NUMBER: US/09/916,359  
; PRIORITY FILING DATE: 2001-07-26  
; PRIOR APPLICATION NUMBER: 09/388,874  
; PRIOR FILING DATE: 1999-09-02  
; PRIOR APPLICATION NUMBER: 97/02,887  
; PRIOR FILING DATE: 1997-03-06  
; NUMBER OF SEQ ID NOS: 2  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 1  
; LENGTH: 9379  
; TYPE: DNA  
; ORGANISM: Virus  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (320)...(9352)  
US-09-916-359-1

Query Match 86.7%; Score 1786; DB 4; Length 9379;  
Best Local Similarity 91.7%; Pred. No. 0;  
Matches 1888; Conservative 0; Mismatches 170; Indels 0; Gaps 0;

QY 2 TGGCGCTATACGGGCTATGCCAGCAGACAAAGGGGCTTTGGGATGCATAATCACCA 61  
DB 3396 TGGCGCCCATACCGGCTACGCCGACGACAAAGGGGCTCTAGGCTGCATAATCACCA 3455  
QY 62 GCTTGACCGCGCGGACAAAACAGGTGAGGTGAGGTGAGGTGAGTCAATGCTCACTG 121  
DB 3456 GCTTAATGCGCGGACAAAACAGGTGAGGTGAGGTGAGGTGAGTCAATGCTCACTG 3515  
QY 122 CCCAGACTTTTGGCAACTGCATTAACCGGGGTGTTGGACTGTCTACCAATGGAGCG 181  
DB 3516 CCCAAACCTTCTGGCAACTGCATTAACCGGGGTGTTGGACTGTCTACCAAGGGCG 3575  
QY 182 GAACAGGACCAATGGCTCACTAAGGTCTGTTATCCAGATGACCAATGTGGACC 241  
DB 3576 GAACAGGACCAATGGCTCACTAAGGTCTGTTATCCAGATGACCAATGTAGACC 3635  
QY 242 AAGACTCTGAGTGGCGCGCTCCCAAGGTGCGGCTCATTAAACCATGCACTTGG 301  
DB 3636 AAGACTCTGAGTGGCGCGCTCCCAAGGTGCGGCTCATTAAACCATGCACTTGG 3695  
QY 302 GCTCTCGGACCTTTACTCTGGTCAAGGACCGCGATGTCATCTGTGCGCGGACGG 361

DB 3696 GCTCTCTCGGACCTTTACTGTGTCAAGAGGACGCGCATGTCTATCCCTGTCGCGCGGG 3755  
QY 362 GTGATGGCAGGGGACGCTGCTTTCCGCCCGGCTATCTCTTACTTTGAAGGCTCTCCGG 421  
DB 3756 GTGATGAGGGGACGCTGCTGCTGCGCGCGGCTATCTCTTACTTTGAAGGCTCTCCGG 3815  
QY 422 GAGGCGCTCTGCTGTGCGCGGACGACATGCCGTAGGACATATTCAGAGCGCGGTATGCA 481  
DB 3816 GGGGTCCGCTGTTGTGCGCGGCGGACGCGGTGGCATATTTAGGGCGCGGTGCA 3875  
QY 482 CCCGTGAGTGGCTAAGCGGTGGACTTCATCTCCCGTAGAGCTTTAGAGCAACATGA 541  
DB 3876 CCCGTGAGTGGCTAAGCGGTGGACTTTATCCCTGTGAGAACTTAGAGCAACATGA 3935  
QY 542 GGTCCCGGCTGTTCTCAGACAACTCTCCCGACAGCAGTCCCGACAGCTACCAAGTGG 601  
DB 3936 GGTCCCGGCTGTTCTCAGAGTAATCTCTCTCCACAGTAGTGTCCCGAGAGCTTCCAGGTGG 3995  
QY 602 CCCACTGCTATGCTCCCGACCGGCTAAGAGCAACCAAGGTCCCGGCGCATACGCA 661  
DB 3996 CTCACCTCCATGCTCCCGACCGGCGGCAAGCAACCAAGGTCCCGGCTGATATGCA 4055  
QY 662 CTCAGGCTACAAGTGTGCTGCTCAACCCCTCCGTTGCTGCAACAATGGGCTTTGGTG 721  
DB 4056 CTCAGGCTACAAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 4115  
QY 722 CTTACATGCTCAAGGCGGCTGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 781  
DB 4116 CTTACATGCTCAAGGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 4175  
QY 782 TCTACGCGAGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 841  
DB 4176 CCACTGGGACCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 4235  
QY 842 CAGGGGTGCTTATGACATAATAATTTGTGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTG 901  
DB 4236 CAGGGGTGCTTATGACATAATAATTTGTGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTG 4295  
QY 902 TCTTGGGCTTGGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 961  
DB 4296 TCTTGGGCTTGGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 4355  
QY 962 TCGCCACCGCTACCCCTCGGCTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1021  
DB 4356 TCGCCACCGCTACCCCTCGGCTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 4415  
QY 1022 CTTCTGCTCACTACCGGAGAGTCCCTTTTATGCAAGGCTATTCCTTGAAGCAATTA 1081  
DB 4416 CTTCTGCTCACTACCGGAGAGTCCCTTTTATGCAAGGCTATTCCTTGAAGCAATTA 4475  
QY 1082 AGGGGGGAGACATCTCATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1141  
DB 4476 AGGGGGGAGACATCTCATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 4535  
QY 1142 AACTGCTCGCTGCGGCTCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1201  
DB 4536 AGCTGCTCGCATTTGGGCTCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 4595  
QY 1202 TCCCGACCAAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1261  
DB 4596 TCCCGACCAAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 4655  
QY 1262 GCGACTTCGATTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1321  
DB 4656 GCGACTTCGATTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 4715  
QY 1322 TTGACCTTACCTTCAACATGAGCAATCAAGCTTCCCGAGGATGCTGTCTCCCGTACTTC 1381  
DB 4716 TTGACCTTACCTTCAACATGAGCAATCAAGCTTCCCGAGGATGCTGTCTCCCGTACTTC 4775  
QY 1382 AACGTCGGGTAGGCTGCGGAGGAAAGCAGGCAATCTACAGATTTGTGGACCGCGGG 1441  
DB 4776 AACGTCGGGTAGGCTGCGGAGGAAAGCAGGCAATCTACAGATTTGTGGACCGCGGG 4835





[illegible]

Db 4618 TCCGACACGCGGATGTTGCTGCTGCAACCGATGCCCTCATGACCGGTATACCG 4677  
QY 1262 GCAGCTTCGATTCGGTGATAGATGCAACACAGTGTGTACCCACAGACGTGCTTCAGCC 1321  
Db 4678 GCAGCTTCGATTCGGTGATAGATGCAACACAGTGTGTACCCACAGACGTGCTTCAGCC 4737  
QY 1322 TTGACCTACCTTCACATTCAGATGACATACAGTGTGTTCACAGGAGTGTGTTCCTCCGCTC 1381  
Db 4738 TTGACCTACCTTCACATTCAGATGACATACAGTGTGTTCACAGGAGTGTGTTCCTCCGCTC 4797  
QY 1382 AAGTCGGGGTACGATTCGACAGGAGGAGCCAGGATCTACAGATTTGTGGACCGGGG 1441  
Db 4798 AAGTCGGGGTACGATTCGACAGGAGGAGCCAGGATCTACAGATTTGTGGACCGGGG 4857  
QY 1442 AGCGTCTCTTCGATGATTTGATCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1501  
Db 4858 AGCGTCTCTTCGATGATTTGATCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 4917  
QY 1502 CTTGGTATGAGCTTACGCGCGCGAGACCAACAGTGTAGGCTACGAGATACATGAAACACCC 1561  
Db 4918 CTTGGTATGAGCTTACGCGCGCGAGACTACAGTGTAGGCTACGAGGCTACATGAAACACCC 4977  
QY 1562 CGGAGCTTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1621  
Db 4978 CGGAGCTTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 5037  
QY 1622 CCACATAGACGCCACTTCTCTTCCAGACCAAGAGTGTGGGAGGAGGCTTCCCTATC 1681  
Db 5038 CTATATAGATGCCACTTCTTATCCAGACCAAGAGTGTGGGAGGAGGCTTCCCTATC 5097  
QY 1682 TGGTAGCGTACCAAGCCCGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1741  
Db 5098 TGGTAGCGTACCAAGCCCGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 5157  
QY 1742 AGATGTGGAAGTCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1801  
Db 5158 AGATGTGGAAGTCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5217  
QY 1802 ATGACTGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1861  
Db 5218 ACAGCTGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 5277  
QY 1862 TGACATGTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1921  
Db 5278 TGACATGTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 5337  
QY 1922 TTGTCGCTGCTTGGCGCGGTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1981  
Db 5338 TCCTGCTGCTTGGCGCGGTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 5397  
QY 1982 TTGCTCTGTCGGAAGCGGCAATCATACCGACAGGAGTCTCTACCGGAGTTCCG 2041  
Db 5398 TCCTGCTGCTTGGCGCGGTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 5457  
QY 2042 ATGAAATGGAAGTGTCT 2059  
Db 5458 ATGAGATGGAGAGTGTCT 5475

RESULT 5  
US-08-440-519-9  
; Sequence 9, Application US/08440519  
; Patent No. 5712087  
; GENERAL INFORMATION:  
; APPLICANT: Houghton, Michael  
; APPLICANT: Choo, Qui-Lim  
; APPLICANT: Kuo, George  
; TITLE OF INVENTION: Combinations of Hepatitis C virus (HCV)  
; NUMBER OF INVENTION: Antigen for Use in Immunoassays for Anti-HCV Antibodies  
; NUMBER OF SEQUENCES: 12  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Chiron Corporation

STREET: P.O. Box 8097 (Int. Prop. R-440)  
CITY: Emeryville  
STATE: CA  
COUNTRY: U.S.A.  
ZIP: 94662-8097  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/440,519  
FILING DATE: 12-MAY-1995  
CLASSIFICATION: 530  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/910,760  
FILING DATE: 07-JUL-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: Blackburn Esq., Robert P.  
REGISTRATION NUMBER: 30,447  
REFERENCE/DOCKET NUMBER: 0101.002  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (510) 601-2702  
TELEFAX: (510) 655-3542  
INFORMATION FOR SEQ ID NO: 9:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 9401 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 342..9374  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: 366  
OTHER INFORMATION: /note= "This amino acid position  
OTHER INFORMATION: can also be Arg."  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: 372  
OTHER INFORMATION: /note= "This amino acid position  
OTHER INFORMATION: can also be Thr."  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: 867  
OTHER INFORMATION: /note= "This amino acid position  
OTHER INFORMATION: can also be Thr."  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: 1341  
OTHER INFORMATION: /note= "This amino acid position  
OTHER INFORMATION: can also be Val."  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: 2148  
OTHER INFORMATION: /note= "This amino acid position  
OTHER INFORMATION: can also be Ile."  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: 2883  
OTHER INFORMATION: /note= "This amino acid position  
OTHER INFORMATION: can also be Asn."  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: 3681  
OTHER INFORMATION: /note= "This amino acid position  
OTHER INFORMATION: can also be Ser."  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: 3690  
OTHER INFORMATION: /note= "This amino acid position

```
;/ OTHER INFORMATION: can also be Thr."/;
;/ FEATURE:;
;/ NAME/KEY: misc_feature;
;/ LOCATION: 4167;
;/ OTHER INFORMATION: /note= "This amino acid position;
;/ OTHER INFORMATION: can also be Leu."/;
;/ FEATURE:;
;/ NAME/KEY: misc_feature;
;/ LOCATION: 4323;
;/ OTHER INFORMATION: /note= "This amino acid position;
;/ OTHER INFORMATION: can also be Val."/;
;/ FEATURE:;
;/ NAME/KEY: misc_feature;
;/ LOCATION: 4701;
;/ OTHER INFORMATION: /note= "This amino acid position;
;/ OTHER INFORMATION: can also be Tyr."/;
;/ FEATURE:;
;/ NAME/KEY: misc_feature;
;/ LOCATION: 4752;
;/ OTHER INFORMATION: /note= "This amino acid position;
;/ OTHER INFORMATION: can also be Ser."/;
;/ FEATURE:;
;/ NAME/KEY: misc_feature;
;/ LOCATION: 5970;
;/ OTHER INFORMATION: /note= "This amino acid position;
;/ OTHER INFORMATION: can also be Gly."/;
;/ FEATURE:;
;/ NAME/KEY: misc_feature;
;/ LOCATION: 6183;
;/ OTHER INFORMATION: /note= "This amino acid position;
;/ OTHER INFORMATION: can also be His."/;
;/ FEATURE:;
;/ NAME/KEY: misc_feature;
;/ LOCATION: 6186;
;/ OTHER INFORMATION: /note= "This amino acid position;
;/ OTHER INFORMATION: can also be Cys."/;
;/ FEATURE:;
;/ NAME/KEY: misc_feature;
;/ LOCATION: 6402;
;/ OTHER INFORMATION: /note= "This amino acid position;
;/ OTHER INFORMATION: can also be Val."/;
;/ FEATURE:;
;/ NAME/KEY: misc_feature;
;/ LOCATION: 7386;
;/ OTHER INFORMATION: /note= "This amino acid position;
;/ OTHER INFORMATION: can also be Ser."/;
;/ FEATURE:;
;/ NAME/KEY: misc_feature;
;/ LOCATION: 7494;
;/ OTHER INFORMATION: /note= "This amino acid position;
;/ OTHER INFORMATION: can also be Phe."/;
;/ FEATURE:;
;/ NAME/KEY: misc_feature;
;/ LOCATION: 7497;
;/ OTHER INFORMATION: /note= "This amino acid position;
;/ OTHER INFORMATION: can also be Ala."/;
;/ FEATURE:;
;/ NAME/KEY: misc_feature;
;/ LOCATION: 7845;
;/ OTHER INFORMATION: /note= "This amino acid position;
;/ OTHER INFORMATION: can also be Phe."/;
;/ FEATURE:;
;/ NAME/KEY: misc_feature;
;/ LOCATION: 8409;
;/ OTHER INFORMATION: /note= "This amino acid position;
;/ OTHER INFORMATION: can also be Gly."/;
;/ FEATURE:;
;/ NAME/KEY: misc_feature;
;/ LOCATION: 9102;
;/ OTHER INFORMATION: /note= "This amino acid position;
;/ OTHER INFORMATION: can also by Gly."/;
;/ FEATURE:;
;/ NAME/KEY: misc_feature;
```

```
;/ LOCATION: 9327;
;/ OTHER INFORMATION: /note= "This amino acid position;
;/ OTHER INFORMATION: can also be Pro."/;
US-08-440-519-9

Query Match      86.7%; Score 1786; DB 1; Length 9401;
Best Local Similarity 91.7%; Pred. No. 0;
Matches 1888; Conservative 0; Mismatches 170; Indels 0; Gaps 0;

QY 2 TGGCGCCTATCAGCGCCTATGCCAGCAGACAGAGGGGGCTTTGGGATGCATAATCACCA 61
Db 3418 TGGCGCCCATCACGGCGTACGCCACAGCAGACAGGGGGCTCCTAGSGTGCAATAACCA 3477

QY 62 GCTTCACCGCGCGGACAAACACAGGTGGAGGTGAGGTTTCAGATCGTGTCAACTGCTG 121
Db 3478 GCCTAACTGGCGCGGACAAACACAGGTGGAGGTGAGGTTTCAGATCGTGTCAACTGCTG 3537

QY 122 CCCAGACTTTCTTGGCAACCTGCATTAAACGGGGTGTGTGGACTGTCTACCATGGAGCG 181
Db 3538 CCCAAACCTTCTCGCAACGTGCATCAATGGGTGTCTGGACTGTCTACCAAGGGCGG 3597

QY 182 GAACAAGGACCATTCGCTACCTTAAGGTCTCTGTATCCAGATGTACACCATGTGGACC 241
Db 3598 GAACGAGGACCATTCGCTACCAAGGTCTCTGTATCCAGATGTATACCAATGTAGACC 3657

QY 242 AAGACCTCGTAGGCTGGCCCGCTCCCAAGTCCCGCTCATTAACACCATGCACCTTGGC 301
Db 3658 AAGACCTTGTGGGCTGGCCCGCTCCGAAGGTAGCCGCTCATTTGAACCCCTGCACTTGGC 3717

QY 302 GCTCCTCGGACCTTTTACCTGTGTACAGGCGACGCGATGTCTCTGTGGCCGACGG 361
Db 3718 GCTCCTCGGACCTTTTACCTGTGTACAGGCGACGCGATGTCTCTCGGTGGCGCGGG 3777

QY 362 GTGATGGCAGGGGCHAGCCTGTCTTTCGCCCGCGGCTATCTCTTACTTTGAAAGCTCTCGG 421
Db 3778 GTGATAGCAGGGGCHAGCCTGTCTTTCGCCCGCGGCTATCTCTTACTTTGAAAGCTCTCGG 3837

QY 422 GAGGCCCTCTGCTGTGCCCGCAGGACATGCCGTAGGCATATTACAGAGCCGGGTATGCA 481
Db 3838 GGGGTCCGCTGTGTGCCCGCGGGGCHAGCCTGTGGGATATTTAGGGCGCGGGTGTGCA 3897

QY 482 CCGGTGGAGTGGCTAAGCGGTGGACTTTTCATCCCGTAGAGAGCTTAGAGCAACATGA 541
Db 3898 CCGGTGGAGTGGCTAAGCGGTGGACTTTTCATCCCGTAGAGAACCTAGAGCAACATGA 3957

QY 542 GGTCCCCGGTGTCTCAGACACTCTCCACACAGAGTGCCTCCAGAGCTTACCAAGTGG 601
Db 3958 GGTCCCCGGTGTCTCAGGATAACTCTCTCCACAGTAGTGCCTCCAGAGCTTCCAGGTGG 4017

QY 602 CCCACTGCTGCTCCACCGGCGAGCGGTAAAGACCAAGGTCCCGGCGCATACGCAG 661
Db 4018 CTCACCTCATGCTCCACAGGCGGCAAAAGCCAAAGGTCCCGGCTGCATATGCAG 4077

QY 662 CTCAGGGCTACAAGGTGTGTGTCTCAACCCCTCCGTTCTGTGCAACAATGGGCTTTGGTG 721
Db 4078 CTCAGGGCTATAAGGTGTGTGTCTCAACCCCTCTGTTGTGCAACACTGGGCTTTGGTG 4137

QY 722 CTTACATGTCCAAGSCCATGGGATTGATCTTAACATCAGGACTGGGGTGAGCAATTA 781
Db 4138 CTTACATGTCCAAGSCTCATGGGATTCGATCTTAACATCAGGACCGGGGTGAGAAATTA 4197

QY 782 CTACTGGCAGCCCGGATCAGTATTCACCTACGGCAAGTTCCTTCCCGACGGCGGTGT 841
Db 4198 CCATCTGGCAGCCCATCAGTACTCCACTACGGCAGTTCCTTCCCGACGGCGGTGT 4257

QY 842 CAGGGGGTCTTATGACATAATAATTTGTGACAGTGCACCTCCACGAGTGAACATCA 901
Db 4258 CGGGGGCGCTTATGACATAATAATTTGTGACAGTGCACCTCCACGAGTGCACATCA 4317

QY 902 TCTTGGGATTTGGCACTGTCTTGTACCAAGCAGACCGCGGGGCGAGACTGACTGTGC 961
Db 4318 TCTTGGGATTCGGCACTGTCTTGTACCAAGCAGAGACTGCGGGGGCGAGACTGGTGTGC 4377
```



LOCATION: 2148  
OTHER INFORMATION: /note= "This amino acid position  
OTHER INFORMATION: can also be Ile."  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: 2883  
OTHER INFORMATION: /note= "This amino acid position  
OTHER INFORMATION: can also be Asn."  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: 3681  
OTHER INFORMATION: /note= "This amino acid position  
OTHER INFORMATION: can also be Ser."  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: 3690  
OTHER INFORMATION: /note= "This amino acid position  
OTHER INFORMATION: can also be Thr."  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: 4167  
OTHER INFORMATION: /note= "This amino acid position  
OTHER INFORMATION: can also be Leu."  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: 4323  
OTHER INFORMATION: /note= "This amino acid position  
OTHER INFORMATION: can also be Val."  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: 4701  
OTHER INFORMATION: /note= "This amino acid position  
OTHER INFORMATION: can also be Tyr."  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: 4752  
OTHER INFORMATION: /note= "This amino acid position  
OTHER INFORMATION: can also be Ser."  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: 5970  
OTHER INFORMATION: /note= "This amino acid position  
OTHER INFORMATION: can also be Gly."  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: 6183  
OTHER INFORMATION: /note= "This amino acid position  
OTHER INFORMATION: can also be His."  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: 6186  
OTHER INFORMATION: /note= "This amino acid position  
OTHER INFORMATION: can also be Cys."  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: 6402  
OTHER INFORMATION: /note= "This amino acid position  
OTHER INFORMATION: can also be Val."  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: 7386  
OTHER INFORMATION: /note= "This amino acid position  
OTHER INFORMATION: can also be Ser."  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: 7494  
OTHER INFORMATION: /note= "This amino acid position  
OTHER INFORMATION: can also be Phe."  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: 7497  
OTHER INFORMATION: /note= "This amino acid position  
OTHER INFORMATION: can also be Ala."

FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: 7845  
OTHER INFORMATION: /note= "This amino acid position  
OTHER INFORMATION: can also be Phe."  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: 8409  
OTHER INFORMATION: /note= "This amino acid position  
OTHER INFORMATION: can also be Gly."  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: 9102  
OTHER INFORMATION: /note= "This amino acid position  
OTHER INFORMATION: can also be Gly."  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: 9327  
OTHER INFORMATION: /note= "This amino acid position  
OTHER INFORMATION: can also be Pro."  
US-08-440-549-9  
Query Match 86.7%; Score 1786; DB 3; Length 9401;  
Best Local Similarity 91.7%; Pred. No. 0;  
Matches 1888; Conservative 0; Mismatches 170; Indels 0; Gaps 0;  
QY 2 TGGCGCCTATCAGCGCCTATGCCAGCAGACAAGGGGCTTTGGGATGCATAATCACA 61  
DB 3418 TGGCGCCCATCAGCGCGTACGCCAGCAGACAAGGGGCTTCTAGGGTGCATAATCACA 3477  
QY 62 GCTTGACCGCGCGGACAAACCAAGGTGGAGGGTGAGGTTTCAGATCGTGCAACTGCTG 121  
DB 3478 GCCTAACTGGCGCGGACAAACCAAGGTGGAGGGTGAGGTTTCAGATCGTGCAACTGCTG 3537  
QY 122 CCCAGACTTCTTGGCAACCTGCATTAACGGGGTGTGTGGACTGTCTACCATGGAGCGG 181  
DB 3538 CCCAAACCTTCTGTGCAACGTGCATCAATGGGGTGTGTGGACTGTCTTACCACGGGGCGG 3597  
QY 182 GAACAAGGACCATTCGCTCAGTCAAGGCTCTGTATCCAGATGTACACCATGTCACCTGCG 241  
DB 3598 GAACGAGGACCATTCGCTCAGTCAAGGCTCTGTATCCAGATGTATACCAATGTAGACC 3657  
QY 242 AAGACCTCTAGGGTGGCGCGCTCCCAAGGTGGCGCTCATTAACACCATGCACTTGG 301  
DB 3658 AAGACCTTGTGGGCTGGCGCGCTCGCAAGGTAGCGGCTCATTTAGACACCTGCACTTGG 3717  
QY 302 GCTCTCGGACCTTTACCTGTGTCAGAGGCGCGCGATGTCTTCTGTGGCGCGGCGG 361  
DB 3718 GCTCTCGGACCTTTACCTGTGTCAGAGGCGCGCGATGTCTTCTGTGGCGCGGCGG 3777  
QY 362 GTGATGGCAGGGGAGCGCTTCTTGGCGCGCGCTATCTTCTTAAAGGCTCCTCGG 421  
DB 3778 GTGATAGCAGGGGAGCGCTTGTGTGCGCGCGCGCAITTTCTTAAAGGCTCCTCGG 3837  
QY 422 GAGGCCCTCTGTGTCGCCCGCAGGACATGCGGTAGGCATATTTCAGAGCGCGGTATGCA 481  
DB 3838 GGGGTCGCTGTGTGTCGCCCGCGGGGACGCGGTGGGATATTATAGGCGCGGGTGTGCA 3897  
QY 482 CCGGTGGAGTGGCTAAGGCGGTGGACTTTCATCCCGGTAGAGAGCTTACAGACAACATGA 541  
DB 3898 CCGGTGGAGTGGCTAAGGCGGTGGACTTTCATCCCTGTGAGAACCTTAGAGACAACATGA 3957  
QY 542 GGTCCCGGCTTCTCAGACAACCTTCCCGCAGCAGTCCCGCAGAGCTACCAAGTGG 601  
DB 3958 GGTCCCGGCTTCTCAGACAACCTTCCCGCAGCAGTCCCGCAGAGCTTCCAGGTGG 4017  
QY 602 CCCACTGCATGCTCCCAACCGGCGGTAAAGACCAAGGTCCTCCCGCGCATACGAG 661  
DB 4018 CTCACCTCCATGCTCCCAACCGGCGGTAAAGACCAAGGTCCTCCCGCATATGAG 4077  
QY 662 CTCAGGGCTACAAGGTGTGTGTGTCACACCTCTCGTTGCTGCAACAATGGGCTTTGGTG 721  
DB 4078 CTCAGGGCTACAAGGTGTGTGTGTCACACCTCTCTGTGCTGCAACAACACTGGGCTTTGGTG 4137

QY 722 CTTTACATGTCCTCAAGCCCATGGGATTTGATCCTTAACATCAGGACTGGGGTGAGSACAATTA 781  
DB 4138 CTTTACATGTCCTCAAGGCTCATGGATCGATCCTTAACATCAGGACCGGGTGAGAACATTA 4197  
QY 782 CTTACTGGCAGCCGATCAGTATTCACCTACGGCAAGTTCTTGGCGAGCGGGTGT 841  
DB 4198 CCACTGGCAGCCCATCAGTATTCACCTACGGCAAGTTCTTGGCGAGCGGGTGT 4257  
QY 842 CAGGGGGTGTATGATGATAAATTTGTGACAGTGCACATCCAGGATGCAACATCCA 901  
DB 4258 CGGGGGGGCTTATGACATTAATTTGTGACAGTGCACATCCAGGATGCAACATCCA 4317  
QY 902 TCTTGGGCAATGTCATGCTCTTGAACAAAGAGAGACCGGGGGGAGACTGACTGTGC 961  
DB 4318 TCTTGGGCAATGTCATGCTCTTGAACAAAGAGAGACTGCGGGGGGAGACTGCTGTGC 4377  
QY 962 TCGCACCCGCTACCCCTCGGGCTCGCTCACTGTGCGCCATCCTTAACATCGAGGAGTTG 1021  
DB 4378 TCGCACCCGCTACCCCTCGGGCTCGCTCACTGTGCGCCATCCTTAACATCGAGGAGTTG 4437  
QY 1022 CTCTGTCCACTACCGGAGAGATCCCTTTTATGGCAAGCTATTCCCTTGAAGCAATTA 1081  
DB 4438 CTCTGTCCACTACCGGAGAGATCCCTTTTATGGCAAGCTATTCCCTTGAAGTHATCA 4497  
QY 1082 AGGGGGGGAGACATCTCATCTTCTGCCACTCAAGGAAGTGCAGAGCTCGCGCAA 1141  
DB 4498 AGGGGGGGAGACATCTCATCTTCTGCCACTCAAGGAAGTGCAGAGCTCGCGCAA 4557  
QY 1142 AAGTGTGCGGTGGGCGTCAATGCGGTGCTTACTACCGGGCTTGTGTGCTCGTCA 1201  
DB 4558 AGTGTGTGCTATGGGATCAATGCGGTGCTTACTACCGGGCTTGTGTGCTCGTCA 4617  
QY 1202 TCCGACCACTGTGAGTGTGCTGTGCACTGCACTGCACTGCACTGCACTGCACTGCA 1261  
DB 4618 TCCGACCACTGTGAGTGTGCTGTGCACTGCACTGCACTGCACTGCACTGCACTGCA 4677  
QY 1262 GCGACTTCGATTCGGTGTAGACTGCAACACGTTGTCAACGAGAGTGCAGCTTCAGCC 1321  
DB 4678 GCGACTTCGATTCGGTGTAGACTGCAACACGTTGTCAACGAGAGTGCAGCTTCAGCC 4737  
QY 1322 TTGACCTACCTTACCATTTAGACAAATCAGCTTCCAGGATGCTGTCTCCGTACTC 1381  
DB 4738 TTGACCTACCTTACCATTTAGACAAATCAGCTTCCAGGATGCTGTCTCCGTACTC 4797  
QY 1382 AAGTCGGGGTGGAGTGCAGAGGAAGCAGGCACTACAGATTTGGGACCGGGG 1441  
DB 4798 AAGTCGGGGTGGAGTGCAGAGGAAGCAGGCACTACAGATTTGGGACCGGGG 4857  
QY 1442 AGCGTCTTCTGTCATGTTGATCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1501  
DB 4858 AGCGCCCTCTCGGATGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 4917  
QY 1502 CTTGGTATGAGTTACGCGCGGAGACCAAGTAGTGTAGGCTACAGATACATGAAACACC 1561  
DB 4918 CTTGGTATGAGTTACGCGCGGAGACCAAGTAGTGTAGGCTACAGATACATGAAACACC 4977  
QY 1562 CGGGACTTCCGCTGTCAGAGCACTTTGAAATTTGGGAGGCGCTTTAGGGTCTCA 1621  
DB 4978 CGGGGCTTCCGCTGTCAGAGCACTTTGAAATTTGGGAGGCGCTTTAGGGGCTCA 5037  
QY 1622 CCACATAGACGCCACTTCTTATCCAGACAAAGCAGAGTGGGAAACCTTCCCTATC 1681  
DB 5038 CTTATATAGTGCCTTCTTATCCAGACAAAGCAGAGTGGGAGACCTTCTTACC 5097  
QY 1682 TGGTAGCTACCAAGCAACGCTGTGCTAGAGCTCAAGCCCTTCCCGTGTGGGAGC 1741  
DB 5098 TGGTAGCTACCAAGCAACGCTGTGCTAGAGCTCAAGCCCTTCCCGTGTGGGAGC 5157  
QY 1742 AGATGTGGAGTGTCTGATCCGCTTCAAGCCCACTTCCAGGCTTCTGCTAT 1801  
DB 5158 AGATGTGGAGTGTCTGATCCGCTTCAAGCCCACTTCCAGGCTTCTGCTAT 5217

QY 1802 ATAGACTGGGCGCTGTCCAGAAATGAAGTCAACCTGACGACCCAGTCAACAAGTATATCA 1861  
DB 5218 ACAGACTGGGCGCTGTTCAGAAATGAAGTCAACCTGACGACCCAGTCAACAATATCA 5277  
QY 1862 TCACATGATGTGCGCTGACCTGAGGTCGTCAAGTACCTGGTGTCTGCTTGGCGG 1921  
DB 5278 TGACATGATGTGCGGCGGACCTGAGGTCGTCAAGGACCTGGGTGCTGCTTGGCGG 5337  
QY 1922 TTCTGGCTGCTTTGGCGCGCTATTGCTTATCCACAGGCTGCTGCTCATAGTAGTAGGA 1981  
DB 5338 TCCTGGCTGCTTTGGCGCGCTATTGCTTCAACAGGCTGCTGCTCATAGTAGTAGGA 5397  
QY 1982 TTGCTTGTTCGGAAGCGGCAATCATACCCGACAGGAAAGTCTCTTACCGGAGTTCG 2041  
DB 5398 TCGTCTTGTTCGGAAGCGGCAATCATACCTGACAGGAAAGTCTCTTACCGAGATTTCG 5457  
QY 2042 ATGAATGGAAGAGTCT 2059  
DB 5458 ATGAGATGGAAGAGTCT 5475

## RESULT 7

US-08-823-895A-25  
; Sequence 25, Application US/08823895A  
; Patent No. 6433159  
; GENERAL INFORMATION:  
; APPLICANT: Kevin P. Anderson  
; TITLE OF INVENTION: Compositions And Methods For  
; TITLE OF INVENTION: Treatment Of Hepatitis C Virus-Associated Diseases  
; NUMBER OF SEQUENCES: 27  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Jane Massey Licata, Esq.  
; STREET: 66 E. Main Street  
; CITY: Marlton  
; STATE: NJ  
; COUNTRY: USA  
; ZIP: 08053

COMPUTER READABLE FORM:  
MEDIUM TYPE: DISKETTE, 3.5 INCH, 1.44 Mb STORAGE  
COMPUTER: IBM 486  
OPERATING SYSTEM: WINDOWS FOR WORKGROUPS  
SOFTWARE: WORDPERFECT 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/823,895A  
FILING DATE: March 17, 1997  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/453,085  
FILING DATE: May 30, 1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/945,289  
FILING DATE: September 10, 1992  
ATTORNEY/AGENT INFORMATION:  
NAME: Jane Massey Licata  
REGISTRATION NUMBER: 32,257  
REFERENCE/DOCKET NUMBER: ISPH-0203  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (609) 779-2400  
TELEFAX: (609) 810-1454  
INFORMATION FOR SEQ ID NO: 25:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 9401  
TYPE: Nucleic  
STRANDEDNESS: Single  
TOPOLOGY: Linear  
ANTI-SENSE: NO

US-08-823-895A-25

Query Match 86.7%; Score 1786; DB 3; Length 9401;

Best Local Similarity 91.7%; Pred. No. 0;

Matches 1888; Conservative 0; Mismatches 170; Indels 0; Gaps 0;

QY 2 TGGCGCTATCAACGCGCTATGCCAGACAGAGGGGCGCTTTTGGGATGCATAATCACCA 61

Db 3418 TGGCGCCATCATCGGCGTACGCCAGCAGACAAGGGCCCTCTAGGTGCAATAACCA 3477  
Qy 62 GCTTGACCGCGCGGAGAAAACACAGGTGGAGGTGAGGTTCAGATCGTGTCAACTGCTG 121  
Db 3478 GCTTAACCTGCGCGGAGAAAACCAAGTGGAGGTGAGGTTCAGATCGTGTCAACTGCTG 3537  
Qy 122 CCAGACTTCTTGGCAACTGCATTAACGGGGTGTGTTGGAAGTGTCTACCATGAGCG 181  
Db 3538 CCACAACTTCTTGGCAAGCTGCATCAATGGGGTGTGCTGGAAGTGTCTACCAAGGGCG 3597  
Qy 182 GAACAGAGCACTTGGCTCACCTAAGGGTCTGTATCCAGATGTACCAATGTGACC 241  
Db 3598 GAACAGAGCACTTGGCTCACCTAAGGGTCTGTATCCAGATGTATACCAATGTAGACC 3657  
Qy 242 AAGACTCGTAGCTGGCGCGCTCCCAAGGTGCGGCTCATTAACACATGACTTGG 301  
Db 3658 AAGACTTGTGGGCTGGCGCGCTCCGCAAGGTAGCGGCTCATTAACACATGACTTGG 3717  
Qy 302 GCTCTCGGACCTTTACCTGGTACGAGGACGCGGATGTCTCTGTGGCGCGGAGGG 361  
Db 3718 GCTCTCGGACCTTTACCTGGTACGAGGACGCGGATGTCTCTGTGGCGCGGAGGG 3777  
Qy 362 GTGATGGAGGGGCAAGCTGCTTTCCGCCCGGCTATCTCTTACTTGAAGGCTCTCG 421  
Db 3778 GTGATAGCAGGGGCAAGCTGCTTTCCGCCCGGCTATCTCTTACTTGAAGGCTCTCG 3837  
Qy 422 GAGGCGCTTGTGTGCGCGCGGAGCATGCGTAGGATATTCAGAGCGCGGATGCA 481  
Db 3838 GGGGTCCGCTGTGTGCGCGCGGAGCATGCGTAGGATATTCAGAGCGCGGATGCA 3897  
Qy 482 CCGCTGGAGTGTCTAAGGGGTGGACTTCCATCCCGTAGAGAGCTTAGAGACAACATGA 541  
Db 3898 CCGCTGGAGTGTCTAAGGGGTGGACTTCCATCCCGTAGAGAGCTTAGAGACAACATGA 3957  
Qy 542 GGTCCCGGTGTCTCAGACAATCTCTCCCAAGCAAGTGTGCGGAGCTTACCAAGTGG 601  
Db 3958 GGTCCCGGTGTCTCAGACAATCTCTCCCAAGCAAGTGTGCGGAGCTTACCAAGTGG 4017  
Qy 602 CCACTGTGATCTCCACCGGAGGGTAAAGCAGCAAGTGTCCGGCGGATACGAG 661  
Db 4018 CTCACCTCATGCTCCACAGGAGCGGCAAGCAAGTGTCCGGCGGATATGCA 4077  
Qy 662 CTCAGGGCTACAAGGTGTGTGTCTCAACCCCTCGTGTCTGCAACAATGGGCTTTGGTG 721  
Db 4078 CTCAGGGCTACAAGGTGTGTGTCTCAACCCCTCGTGTCTGCAACAATGGGCTTTGGTG 4137  
Qy 722 CTTCATGTCCAAAGGCCATGGATTTGATCTTAACATCAGGACTGGGGTGAGGACAATTA 781  
Db 4138 CTTCATGTCCAAAGGCCATGGATTTGATCTTAACATCAGGACTGGGGTGAGGACAATTA 4197  
Qy 782 CTACTGGGAGCGGATCAGTATTCACCTAGCGCAAGTTCCTTGGCGGAGGGGTGT 841  
Db 4198 CCACTGGGAGCGGATCAGTATTCACCTAGCGCAAGTTCCTTGGCGGAGGGGTGT 4257  
Qy 842 CAGGGGTCTTATGACATAATAATTTGAGCAGTGGCCACTCCACGAGTCAACATCCA 901  
Db 4258 CAGGGGTCTTATGACATAATAATTTGAGCAGTGGCCACTCCACGAGTCAACATCCA 4317  
Qy 902 TCTTGGGCATTTGGCACTGCTTTGACCAAGAGAGACCGCGGGGCGAGACTGACTGTGC 961  
Db 4318 TCTTGGGCATTTGGCACTGCTTTGACCAAGAGAGACTGCGGGGCGAGACTGGTGTGC 4377  
Qy 962 TCGCCACCGCTACCCCTCGGGCTCGGTCACTGTGCCCATCTTAACATCGAGGAGGTG 1021  
Db 4378 TCGCCACCGCTACCCCTCGGGCTCGGTCACTGTGCCCATCTTAACATCGAGGAGGTG 4437  
Qy 1022 CTCTGTCCACTACCGGAGAGATCCCTTTTATGGCAAGGCTATTCCTCTTGAAGCAATTA 1081  
Db 4438 CTCTGTCCACTACCGGAGAGATCCCTTTTATGGCAAGGCTATTCCTCTTGAAGTAAATCA 4497  
Qy 1082 AGGGGGGAGACATCTCTCTCTGCGCACTCAAGAGAGTGTGCAAGGCTCGCGGCA 1141

Db 4498 AGGGGGGAGACATCTCATCTTCTGTCTATTCAAAGAGAAGTGTGCAAGAACTCGCGCAA 4557  
Qy 1142 AACTGGTTCGGTGGCGCTCAATGCGCGGTCTTACTACGGCGCTTGTATGTGTCGCTCA 1201  
Db 4558 AGCTGGTCGATTTGGGCATCAATGCGCGGTCTTACTACGGCGCTTGTGACGTTGCGTCA 4617  
Qy 1202 TCCCGACAGTGTGACGCTTGTGCTGCTGGCAACTGAGCGCCCTCATGACCGGCTTTACCG 1261  
Db 4618 TCCCGACAGCGCGCATGTTGCTGCTGGTGGCAAGCCGATGCCCTCATGACCGGCTATACG 4677  
Qy 1262 GCGACTTCGATTTGGTGTATAGCTGCAACACGTTGTGTCAACAGACAGTGTGACTTCAGCC 1321  
Db 4678 GCGACTTCGATTTGGTGTATAGCTGCAACACGTTGTGTCAACAGACAGTGTGACTTCAGCC 4737  
Qy 1322 TTGACCTTACCTTACCATTTGAGCAATCAGCTTCCCGAGGATGCTCTCCCGTACTC 1381  
Db 4738 TTGACCTTACCTTACCATTTGAGCAATCAGCTTCCCGAGGATGCTCTCCCGTACTC 4797  
Qy 1382 AACGTCGGGGTAGGACTGGCAGAGGGAAGCAGGATCTTACAGATTTGTGGCACCGGGG 1441  
Db 4798 AACGTCGGGGTAGGACTGGCAGAGGGAAGCAGGATCTTACAGATTTGTGGCACCGGGG 4857  
Qy 1442 AGCGTCTTCTGGCATTTGTGATGCTGCTCTCTGCGAGTGTGTATGACGGGGTGTG 1501  
Db 4858 AGCGTCTTCTGGCATTTGTGATGCTGCTCTCTGCGAGTGTGTATGACGGGGTGTG 4917  
Qy 1502 CTTGTGTAGCTTACGCGCGGAGACCAAGTGTAGGCTAGGACATACATGACATGACACCC 1561  
Db 4918 CTTGTGTAGCTTACGCGCGGAGACCAAGTGTAGGCTAGGACATACATGACATGACACCC 4977  
Qy 1562 CGGACTTCCCGTGTGCGCAAGACCATCTTGAATTTTGGGAGGGCGTCTTTACGGTCTCA 1621  
Db 4978 CGGACTTCCCGTGTGCGCAAGACCATCTTGAATTTTGGGAGGGCGTCTTTACAGGCTCA 5037  
Qy 1622 CCCACATAGACGCCCATCTTCTTATCCAGACAAGAGAGTGGGGAACCTTCCCTATC 1681  
Db 5038 CTCATATAGATGCCCATCTTCTTATCCAGACAAGAGAGTGGGGAACCTTCCCTATC 5097  
Qy 1682 TGGTAGCTTCCAGAGCCACCGTGTGCTAGAGCTCAAGCCCTCCCGTGTGGGAGCC 1741  
Db 5098 TGGTAGCTTCCAGAGCCACCGTGTGCTAGAGCTCAAGCCCTCCCGTGTGGGAGCC 5157  
Qy 1742 AGATGGAAGTGTGTGATCGCTTCAAGCCACCCCTCATGGGCCAACACCTCTGCTAT 1801  
Db 5158 AGATGGAAGTGTGTGATCGCTTCAAGCCACCCCTCATGGGCCAACACCTCTGCTAT 5217  
Qy 1802 ATAGACTGGGCGCTGTCCAGAAATGAAGTCAAGCTGACGACCCAGTCAACAGTATATCA 1861  
Db 5218 ACAGACTGGGCGCTGTTCAGAAATGAAGTCAAGCTGACGACCCAGTCAACAAATATCA 5277  
Qy 1862 TGACATGTATGTGGGCTGACCTGGAGGTGCTCAGAGTACCTGGGTGCTCGTTGGCGG 1921  
Db 5278 TGACATGTATGTGGGCGGACCTGGAGGTGCTCAGAGCACTGGGTGCTCGTTGGCGG 5337  
Qy 1922 TTTGGCTGTCTTGGCGCGCTATTGCTTATCCAGAGGTGCTGCTGCTATAGTAGGAGGA 1981  
Db 5338 TCTTGGCTGTCTTGGCGCGCTATTGCTTATCCAGAGGTGCTGCTGCTATAGTAGGAGGA 5397  
Qy 1982 TTTGTCTTGTCCGGAAGCGCGCAATCATACCGGACAGGGAAGTCTCTTACCGGGAGTTCG 2041  
Db 5398 TCGTCTTGTCCGGAAGCGCGCAATCATACCTGACAGGGAAGTCTCTTACCGAGAGTTCG 5457  
Qy 2042 ATGAAATGGAAGTGTCT 2059  
Db 5458 ATGAGATGGAAGTGTCT 5475

RESULT 8  
PCT-US91-02225-9  
; Sequence 9, Application PC/TUS9102225  
; GENERAL INFORMATION:  
; APPLICANT: HOUGHTON, MICHAEL  
; APPLICANT: CHOO, QUI-LIM





Qy	1562	CGGGA	CTTCCCGTGTG	CCAAAGAC	ATCTTGAA	TTTTGGAGGGCGCTTTTA	CGGCTCTCA	1621
Db	4978	CGGGGCT	TCCCGTGTG	CCAGGAC	CACTTGGAA	TTTTGGAGGGCGCTTTTAC	AGGGCTCA	5037
Qy	1622	CCACATAGAC	GGCCCACTT	CCATCC	CAGACAAG	CAGAGTGGGGA	AAACCTTTCCCTATC	1681
Db	5038	CTCATATAGAT	GCCCACTT	CTATCC	CAGACAAG	CAGAGTGGGGA	GAACCTTTCCCTTACC	5097
Qy	1682	TGTTAGCGTA	CCAAAGCACA	CCGTGTG	CGCTAGAGCT	CAAGCCCTCTCCCGCTCGTGGGACC	1741	
Db	5098	TGTTAGCGTAT	CCAAAGCACC	CGTGTG	CGCTAGAGGCT	CAAGCCCTCTCCCGCATCGTGGGACC	5157	
Qy	1742	AGATGTGGAA	GTGCTTGAT	CCGTCTCA	AGGCCCA	CCCTCCATGGGCCCA	ACACCTTCTGCTAT	1801
Db	5158	AGATGTGGA	GTGTTGAT	TGCGCTCA	AGGCCCA	CCCTCCATGGGCCCA	ACACCTTCTGCTAT	5217
Qy	1802	ATGACTGGGG	CGGCTGTCC	AGAAATGA	AGTACACCT	TGACCGCACCCAGTCA	CCAAGTATATCA	1861
Db	5218	ACAGACTTGGG	CGCTGTTT	CAGAAATGA	ATCACCTCTG	ACGCAACCCAGTCA	CCAAATATCATCA	5277
Qy	1862	TGACATGTAT	TGTCGGCTG	ACTGGAGGT	CGTAC	CGAGTACCTGGGTGCTCGTTGGGGGGG	1921	
Db	5278	TGACATGCAT	TGTCGGCCGA	CTTGGAGGT	CGTCA	CGAGCACCTGGGTGCTCGTTGGGGGGG	5337	
Qy	1922	TTCTGGCTGTCT	TTGGCCGGGTA	TTGCGCTAT	TCAC	CAGGCTCGGTGGTCA	TAGTAGGTAGGA	1981
Db	5338	TCCTGGCTGTCT	TTTGGCCGGCTAT	TGCGCTGT	CAAC	CAGGCTCGGTGGTCA	TAGTGGGACGGG	5397
Qy	1982	TTGTCTTTGT	CCGGAAAC	CCGGCAAT	CATAC	CCCGACAGGGAAGT	CCTCTACCCGGGAGTTCCG	2041
Db	5398	TCGHTCTTGT	CCGGAGAC	CCGGCAAT	CATAC	CCTGACAGGGAAGT	CCTCTACCGAGAGTTCCG	5457
Qy	2042	ATGAAATGGA	AGAGTGTCT	2059				
Db	5458	ATGAGATGGA	AGAGTGTCT	5475				

## RESULT 9

US-08-444-818-85  
Sequence 65, Application US/08444818  
Patent No. 6150087  
GENERAL INFORMATION:  
APPLICANT: Chien, David Y.  
APPLICANT: Rutter, William J.  
TITLE OF INVENTION: NANBV Diagnostics and Vaccines  
NUMBER OF SEQUENCES: 777  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Chiron Corporation  
STREET: 4560 Horton Street  
CITY: Emeryville  
STATE: CA  
COUNTRY: USA  
ZIP: 94608-2916  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
FILING DATE: US/08/444,818  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/403,590  
FILING DATE: 14-MAR-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Harbin, Alisa A.  
REGISTRATION NUMBER: 33,895  
REFERENCE/DOCKET NUMBER: 0110.002  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (508)359-3876  
TELEFAX: (508)359-3885  
INFORMATION FOR SEQ ID NO: 65:



QY	182	GAACAAAGGACCAATGGCTCACTAAGGGTCTGTATTCAGATGTACACCAATGTGGACC	241	Db	3995	GCAGTTCGACTCGGTGATAGACTGCAATACAGTGTGTCAACCCAGACAGTCCGATTCAGCC	4054
Db	2915	GAACGAGGACCAATCGGTCACCAAGGGTCTGTATTCAGATGTATACCAATGTAGACC	2974	QY	1322	TTGACCCCTACCTTCAACATTTGACACATACAGCTTCCGAGGATGTGTCTCCCGTACTC	1381
QY	242	AAGACCTCTAGGCTGGCCGCTCCCAAGGTGCCGCTCATTTACACACCATGCACTTGGC	301	Db	4055	TTGACCCCTACCTTCAACATTTGACACATACAGCTTCCGAGGATGTGTCTCCCGTACTC	4114
Db	2975	AAGACCTTGTGGCTGGCCGCTCCCAAGGTAGCGCTCATTTGACACCTTGCATTTGGC	3034	QY	1382	AACGTCCGGGTAGGACTGGCAGAGGAGGACAGGCATCTACAGATTTGTGGCACCGGGGG	1441
QY	302	GCTCTCTCGGACCTTTACCTTGTGTCAGGAGCAGCCGATGTCTTCTGTGGCCGACGGG	361	Db	4115	AACGTCCGGGAGGACTGGCAGAGGAGGACAGGCATCTACAGATTTGTGGCACCGGGGG	4174
Db	3035	GCTCTCTCGGACCTTTACCTTGTGTCAGGAGCAGCCGATGTCTTCTGTGGCCGACGGG	3094	QY	1442	AGCGTCTCTTCTGGCATGTTTGAATCTGTCTCTGTGGAGTCTATGACCGGGGTGTG	1501
QY	362	GTGATGCGAGGGCAGCTGCTTTGCGCCCGGCTATCTTACTTTGAAAGCTCCTCGG	421	Db	4175	AGCGCCCTCTCGGCATGTTTGAATCTGTCTCTGTGGAGTCTATGACCGGGGTGTG	4234
Db	3095	GTGATGCGAGGGCAGCTGCTTTGCGCCCGGCTATCTTACTTTGAAAGCTCCTCGG	3154	QY	1502	CTTGATGAGCTTACCGCCCGGACACAGTACAGTACAGCATACATGAACACCC	1561
QY	422	GAGGCCCTCTGCTGTCGCCCGAGGACATGCCGTAGGCATATTCAGAGCCGCGGTATGCA	481	Db	4235	CTTGATGAGCTTACCGCCCGGACACAGTACAGTACAGCATACATGAACACCC	4294
Db	3155	GGGGTCCGCTGTTGTGCCCCCGGGCAGCGCTGGGCATATTTAGGGCCGCGGTGCA	3214	QY	1562	CGGACTTCCCGTGTGCAAGACCATCTTGAAATTTTGGGAGGGCGTCTTTACCGGTCTCA	1621
QY	482	CCGCTGGAGTGGCTAAGCGGTGGACTTCATCCCGTGTAGAGCTTTAGAGACAAACATGA	541	Db	4295	CGGGGCTTCCCGTGTGCCAGGACCATCTTGAAATTTTGGGAGGGCGTCTTTACAGGCCTCA	4354
Db	3215	CCGCTGGAGTGGCTAAGCGGTGGACTTCATCCCGTGTAGAGCTTTAGAGACAAACATGA	3274	QY	1622	CCCAATAGAGCGCCACTTCTTATCCAGACAAAGAGAGTGGGAAAACCTTCCCTATC	1681
QY	542	GGTCCCGGTGTTCTCAGACAACTCTCCCAAGCAGCTGCCCCAGAGCTACCAAGTGG	601	Db	4355	CTCATATAGATGCCACTTCTTATCCAGACAAAGAGAGTGGGAGAACCTTCTCTTACC	4414
Db	3275	GGTCCCGGTGTTCTCAGAGTAACTCTCTCTCCAGTAGTGGCCAGAGCTTCCAGGTGG	3334	QY	1682	TGGTAGGCTACCAAGCCAGCTGTGCGTCAAGCCCTTCCCGCTCGTGGGACC	1741
QY	602	CCCACTGCACTGCTCCCAAGCGGTAAAGAGCACCAGGCTCCCGCCGATACGCGAG	661	Db	4415	TGGTAGGCTACCAAGCCAGCTGTGCGTCAAGCCCTTCCCGCTCGTGGGACC	4474
Db	3335	CTACCTCCATGCTCCCAAGCGGTAAAGAGCACCAGGCTCCCGCCGATACGCGAG	3394	QY	1742	AGATGTGGAAGTGTGATCCGCTCAAGCCCAACCTCTCATGGGCCAACCTCTGCTAT	1801
QY	662	CTAGGGCTAAGAGTGTGGTGTCAACCCCTCGTGTGTGCAACAAATGGGCTTTGGTG	721	Db	4475	AGATGTGGAAGTGTGATCCGCTCAAGCCCAACCTCTCATGGGCCAACCTCTGCTAT	4534
Db	3395	CTAGGGCTAAGAGTGTGGTGTCAACCCCTCGTGTGTGCAACAAATGGGCTTTGGTG	3454	QY	1802	ATAGCTGGGCGCTGTCCAGATGAGTCAACCTCGAGCCAGCCAGTCCCAAGTATATCA	1861
QY	722	CTTACATGCTCAAGGCTCATGGATCGATCTTACATCAGAGCTGGGGTGGAGCAATTA	781	Db	4535	ACAGACTGGGCGCTGTTCAGAAATGAAATCACTTCAGCCAGCCAGTCCACCAATATCA	4594
Db	3455	CTTACATGCTCAAGGCTCATGGATCGATCTTACATCAGAGCTGGGGTGGAGCAATTA	3514	QY	1862	TGACATGATGTCTGGCTGACCTGGAGGTCTGTCAGAGTACCTGGGTGTCTGTCGGGCG	1921
QY	782	CTACTGGCAGCCGATACAGTATTTCCACCTACGGCAAGTTCTTTCGCGAGCGGGGTGT	841	Db	4595	TGACATGATGTCTGGCTGACCTGGAGGTCTGTCAGAGTACCTGGGTGTCTGTCGGGCG	4654
Db	3515	CCACTGGCAGCCGATACAGTATTTCCACCTACGGCAAGTTCTTTCGCGAGCGGGGTGT	3574	QY	1922	TTCTGGCTGCTTTGGCCGCTATTCCTCATCCAGAGTGTGGTGTCTAGTAGGTAGGA	1981
QY	842	CAGGGGTGTCTATGACATATATTTGTGACAGTGCACCTCCAGGATGCAACATCCA	901	Db	4655	TCCTGGCTGCTTTGGCCGCTATTCCTCATCCAGAGTGTGGTGTCTAGTAGGTAGGA	4714
Db	3575	CAGGGGTGTCTATGACATATATTTGTGACAGTGCACCTCCAGGATGCAACATCCA	3634	QY	1982	TTGTCTTGTTCGGAAGCCGCAATCATACCCGACAGGAAAGTCTCTACCGGGAGTTCCG	2041
QY	902	TCCTGGCTGCTTTGGCCGCTATTCCTCATCCAGAGTGTGGTGTCTAGTAGGTAGGA	961	Db	4715	TCGTCTTGTTCGGAAGCCGCAATCATCTGACAGGGAAGTCTCTCTACCGAGAGTTCCG	4774
Db	3635	TCCTGGCTGCTTTGGCCGCTATTCCTCATCCAGAGTGTGGTGTCTAGTAGGTAGGA	3694	QY	2042	ATGAAATGGAAGAGTGTCT 2059	
QY	962	TCGCGACCGTACCCCTCGGGCTCGTCACTGTGCCCCATCTTAAATCATCGAGAGTTG	1021	Db	4775	ATGAGTGAAGAGTGTCT 4792	
Db	3695	TCGCGACCGTACCCCTCGGGCTCGTCACTGTGCCCCATCTTAAATCATCGAGAGTTG	3754				
QY	1022	CTCTGTCCACTACCGGAGATCCCTTTTATGCAAGGCTATTTCCCTTTGAAGCAATTA	1081				
Db	3755	CTCTGTCCACTACCGGAGATCCCTTTTATGCAAGGCTATTTCCCTTTGAAGTATCA	3814				
QY	1082	AGGGGGGAGACATCTCATCTTCTGCCACTCAAAAGAAAGTGGCAGAGCTCGCCGCAA	1141				
Db	3815	AGGGGGGAGACATCTCATCTTCTGCCACTCAAAAGAAAGTGGCAGAGCTCGCCGCAA	3874				
QY	1142	AACGTGCTGGGTGAGTCAATGCGGTGCTTACTACCGGGCTGTGATGTGCTGCTCA	1201				
Db	3875	AGCTGGTCCGATTTGGGATCAATGCCGTGGCTTACTACCGGGCTTTGAGGTGCTGCTCA	3934				
QY	1202	TCCCGACAGTGTGACAGTGTGCTGCTGGCAATGACCGCCCTCATGACCGGCTTACCG	1261				
Db	3935	TCCCGACAGTGTGACAGTGTGCTGCTGGCAATGACCGCCCTCATGACCGGCTTACCG	3994				
QY	1262	CGGACTTCGATTCGGGTGATAGCTGCAACACGTGTGTCAACCCAGACAGTGTACGCC	1321				

RESULT 11  
US-08-444-818-137  
; Sequence 137, Application US/08444818  
; Patent No. 6150087  
; GENERAL INFORMATION:  
; APPLICANT: Chien, David Y.  
; APPLICANT: Rutter, William J.  
; TITLE OF INVENTION: NANBV Diagnostics and Vaccines  
; NUMBER OF SEQUENCES: 777  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Chiron Corporation  
; STREET: 4560 Horton Street  
; CITY: Emeryville  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 94608-2916  
; COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/444,818  
FILING DATE:  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/403,590  
FILING DATE: 14-MAR-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Harbin, Alisa A.  
REGISTRATION NUMBER: 33,895  
REFERENCE/DOCKET NUMBER: 0110.002  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (508)359-3876  
TELEFAX: (508)359-3885  
INFORMATION FOR SEQ ID NO: 137:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 9987 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 1..9985

US-08-444-818-137

Query Match 86.6%; Score 1784.4; DB 3; Length 9987;  
Best Local Similarity 91.7%; Pred. No. 0;  
Matches 1887; Conservative 0; Mismatches 171; Indels 0; Gaps 0;

QY	2	TGGCGCTATCAGCGGCTATGCCAGCAGACAGAGGGCGCTTTGGGATGCATTAATCACCA	61
DB	3077	TGGCGGCCATCAGCGGGGTACGCCAGCAGACAGAGGGCGCTTCCAGGGTGATTAATCACCA	3136
QY	62	GCTTTAGCCGCGCGGACAAAACACAGGTGAGGGTTCAGATCGTGTCAACTGCTG	121
DB	3137	GCCTAAGTGGCGGACAAAACCAAGTGGGGTGGGTCAGATGTGTCAACTGCTG	3196
QY	122	CCAGACTTTCTGGCAACTGATTAACCGGGGTGTGGGATGTGTCAACTGGAGCCG	181
DB	3197	CCCAAACTTTCTGGCAACTGATTAACCGGGGTGTGGGATGTGTCAACTGGAGCCG	3256
QY	182	GACACAGGACCATGGTGCACCTAAGGGTCTGTATCCAGATGTACCAATGTGGAC	241
DB	3257	GAACAGGACCATGGGTCACCAAGGGTCTGTATCCAGATGTATACCAATGTAGAC	3316
QY	242	AAGACCTCGTAGGTGGCGCGCTCCCAAGGTGCCGCTCATTAACACCATGCACTTGG	301
DB	3317	AAGACCTTGTGGCTGGCGCGCTCCCAAGGTAGCGCTCATTAACACCATGCACTTGG	3376
QY	302	GCTCTCGGACCTTTACCTGGTCAAGGACACCGCGATGTATCTCTGTGCGCCGCGG	361
DB	3377	GCTCTCGGACCTTTACCTGGTCAAGGACACCGCGATGTATCTCTGTGCGCGCGG	3436
QY	362	GTGATGCGGCGGACGCTGCTTCCGCCCGGCTATCTTACTTGAAGGCTCTCTCG	421
DB	3437	GTGATGCGGCGGACGCTGCTTCCGCCCGGCTATCTTACTTGAAGGCTCTCTCG	3496
QY	422	GAGGCCCTCTGCTGTGCCCGGACGACATCCGCTAGGCATATTACAGAGCGCGGTATGCA	481
DB	3497	GAGGCCCTCTGCTGTGTGCCCGGACGACATCCGCTAGGCATATTAGGCGCGCGGTGCA	3556
QY	482	CCCGTGGAGTGGCTAAGGCGGTGGACTTCAATCCCGGTAGAGCTTAGAGACAACCATGA	541
DB	3557	CCCGTGGAGTGGCTAAGGCGGTGGACTTCAATCCCGGTAGAGCTTAGAGACAACCATGA	3616
QY	542	GGTCCCCGGTGTCTCAGACAATCTCTCCCAACAGAGTGGCCCGAGAGTCAACAGTGG	601
DB	3617	GGTCCCCGGTGTCTCAGACAATCTCTCTCCCAACAGAGTGGCCCGAGAGTCTCCAGTGG	3676

QY	602	CCCACCTGCATGCTCCACCGGAGCGGTAAGAGCACCAAGGTCCCGGCGCATACGCAG	661
DB	3677	CTCACCTCCATGCTCCACCGGAGCGGCAAAAGCACCAAGGTCCCGGCTGCATATGCAG	3736
QY	662	CTCAGGGCTACAAGGTGCTGCTCAACCCCTCCGTTGCTGCAACAATGGGCTTTGGTG	721
DB	3737	CTCAGGGCTATAAGGTGCTAGTACTCAACCCCTCTGTTGCTGCAACACTGGGCTTTGGTG	3796
QY	722	CTTACATGTCGAAGGCCCATGGATTGATCTTAAATCAGGACTGGGGTGAGACAATTA	781
DB	3797	CTTACATGTCGAAGGCTCATGGATCGATCTTAAATCAGGACCGGGTGAGACAATTA	3856
QY	782	CTACTGGCAGCCCGATCACGTATTCCACCTACGGCAAGTTCTTTCGCGAGCGGGTGT	841
DB	3857	CCACTGGCAGCCCGATCACGTATTCCACCTACGGCAAGTTCTTTCGCGAGCGGGTGT	3916
QY	842	CAGGGGCTCTTATGACATAAATTTGTGACAGGTGCCACTCCACGGATGCAACATCCA	901
DB	3917	CGGGGGCGCTTATGACATAAATTTGTGACAGGTGCCACTCCACGGATGCAACATCCA	3976
QY	902	TCTTGGGCATTTGGCACTGCTTACCAAGCAGAGACCGGGGGGAGACTCAGCTGTGC	961
DB	3977	TCTTGGGCATTTGGCACTGCTTACCAAGCAGAGACTCGGGGGGAGACTCGTGTGC	4036
QY	962	TCGCCACCGCTACCCCTCCGGGCTCCGTCATCTGTGCCCATCTTAAATCGAGAGGTTG	1021
DB	4037	TCGCCACCGCTACCCCTCCGGGCTCCGTCATCTGTGCCCATCTTAAATCGAGAGGTTG	4096
QY	1022	CTCTGTCACTACCGGAGAGATCCCTTTTATGGCAAGGCTATTCCCTTTGAAGCAATTA	1081
DB	4097	CTCTGTCACTACCGGAGAGATCCCTTTTATGGCAAGGCTATTCCCTTTGAAGCAATTA	4156
QY	1082	AGGGGGGAGAGACTCTCATCTTCTGCACTCAAAAGAAAGTGCAGAGCTCGCGCAA	1141
DB	4157	AGGGGGGAGAGACTCTCATCTTCTGCACTCAAAAGAAAGTGCAGAGCTCGCGCAA	4216
QY	1142	AACTGTCTCGCTTGGGCGCTCAATGCGGCTTACTTACCGGCGCTTGTGTCTCGTCA	1201
DB	4217	AGTGTCTCGCTTGGGCGCTCAATGCGGCTTACTTACCGGCGCTTGTGTCTCGTCA	4276
QY	1202	TCCGACAGTGGTGTGCTGTGCTGGGCAACGACGCTCTCATGACGGGTTTACCG	1261
DB	4277	TCCGACAGTGGTGTGCTGTGCTGGGCAACGACGCTCTCATGACGGGTTTACCG	4336
QY	1262	GGGACTTCCGCTTGGGCTGTAGTCAACACGCTGTGTACCCAGACAGCTTCAGCC	1321
DB	4337	GGGACTTCCGCTTGGGCTGTAGTCAACACGCTGTGTACCCAGACAGCTTCAGCC	4396
QY	1322	TTGACCTTACCTTACCATTTAGACAATCACCGCTTCCCGAGATGTCTCTCCGCTATC	1381
DB	4397	TTGACCTTACCTTACCATTTAGACAATCACCGCTTCCCGAGATGTCTCTCCGCTATC	4456
QY	1382	AACGTCGGGCTAGGACTGGCAGAGGAGCGAGCATCTACAGATTTGTGGACCGGGGG	1441
DB	4457	AACGTCGGGCTAGGACTGGCAGAGGAGCGAGCATCTACAGATTTGTGGACCGGGGG	4516
QY	1442	AGCGTCTTCTGCAATGTTGACTCGTCTGCTGCGAGTGTATGACGCGGGTGTG	1501
DB	4517	AGCGTCTTCTGCAATGTTGACTCGTCTGCTGCGAGTGTATGACGCGGGTGTG	4576
QY	1502	CTTGTATGAGCTTACCGCGCGGAGACCACTAGGCTACGAGATCATGAACACCC	1561
DB	4577	CTTGTATGAGCTTACCGCGCGGAGACCACTAGGCTACGAGATCATGAACACCC	4636
QY	1562	CGGGAATTCCTCGGTGCGCAAGACCATCTTTGAATTTTGGAGGGCGCTTTACCGGCTCA	1621
DB	4637	CGGGAATTCCTCGGTGCGCAAGACCATCTTTGAATTTTGGAGGGCGCTTTACCGGCTCA	4696
QY	1622	CCCACATAGACCGCCACTTCTTATCCAGACAAGCAGAGTGGGGAACCTTCCCTATC	1681
DB	4697	CTCATATAGTCCCACTTCTTATCCAGACAAGCAGAGTGGGGAACCTTCTTATC	4756

QY 1682 TGTAGCTACCAAGACCGCTGTGGCTAGAGCTCAAGCCCTCCCGCTCGTGGGACC 1741  
Db TGTAGCTACCAAGACCGCTGTGGCTAGAGCTCAAGCCCTCCCGCTCGTGGGACC 4816  
QY 1742 AGATGTGGAAGTGTGTGATCCGCTTCAAGCCCACTCCATCGGCGCAACACCTCTGTGCTAT 1801  
Db AGATGTGGAAGTGTGTGATCCGCTTCAAGCCCACTCCATCGGCGCAACACCTCTGTGCTAT 4876  
QY 1802 ATAGACTGGCGGCTGTCCAGATGAGTCACTGAGTACCTGAGCAGCCAGTCAACAGTATATCA 1861  
Db ACAGACTGGCGGCTGTTCAGATGAGTCACTGAGCAGCCAGTCAACAGTATATCA 4936  
QY 1862 TGACATGTATGTCCGCTGACCTGGAGTGTCTCAAGTACCTGGGTGCTGCTTGGCGGG 1921  
Db TGACATGTATGTCCGCTGACCTGGAGTGTCTCAAGTACCTGGGTGCTGCTTGGCGGG 4996  
QY 1922 TTCTGTGCTGTTGGCGGCTATTTGCTATTCACAGGCTGCGTGTGCTATGAGTAGGA 1981  
Db TTCTGTGCTGTTGGCGGCTATTTGCTATTCACAGGCTGCGTGTGCTATGAGTAGGA 4997  
QY 1982 TTGCTTGTCCGGAAGCGGCAATCATACCCGACAGGGAAGTCTCTTACCGGAGTTCG 2041  
Db TTGCTTGTCCGGAAGCGGCAATCATACCCGACAGGGAAGTCTCTTACCGGAGTTCG 5056  
QY 2042 ATCAATGGAAGTGTCT 2059  
Db ATCAATGGAAGTGTCT 5134

RESULT 12  
US-08-444-818-122  
; Sequence 122, Application US/08444818  
; Patent No. 6150087  
; GENERAL INFORMATION:  
; APPLICANT: Chien, David Y.  
; APPLICANT: Rutter, William J.  
; TITLE OF INVENTION: NANBV Diagnostics and Vaccines  
; NUMBER OF SEQUENCES: 777  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Chiron Corporation  
; STREET: 4560 Horton Street  
; CITY: Emeryville  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 94608-2916  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/444,818  
; FILING DATE:  
; CLASSIFICATION: 424  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/08/403,590  
; FILING DATE: 14-MAR-1995  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Harbin, Alisa A.  
; REGISTRATION NUMBER: 33,895  
; REFERENCE/DOCKET NUMBER: 0110,002  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (508)359-3876  
; TELEFAX: (508)359-3885  
; INFORMATION FOR SEQ ID NO: 122:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 9185 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: cDNA  
US-08-444-818-122

Query Match 86.6%; Score 1784.4; DB 3; Length 9185;  
Best Local Similarity 91.7%; Pred. No. 0;  
Matches 1887; Conservative 0; Mismatches 171; Indels 0; Gaps 0;  
QY 2 TGGCGCCCTATCAACGCGCTATGCCAGCAGACAAAGGGGCTTTTGGATGCAATAACCA 61  
Db TGGCGCCCTATCAACGCGCTATGCCAGCAGACAAAGGGGCTTTTGGATGCAATAACCA 3396  
QY 62 GCTTACACCGCGCGGACAAAACACAGGTGAGGTGAGGTTCAGATCGTCACTGCTG 121  
Db GCTTACACCGCGCGGACAAAACACAGGTGAGGTGAGGTTCAGATCGTCACTGCTG 3456  
QY 122 CCCAGACTTTCTTTGGCAACCTGATTAACGGGGTGTGTGGACTGTCTACCAATGGAGCG 181  
Db CCCAGACTTTCTTTGGCAACCTGATTAACGGGGTGTGTGGACTGTCTACCAATGGAGCG 3575  
QY 182 GAACAAGGACCAATTCGCTCACTAAGGTCTCTTTATCCAGATGTACCAATGTGGACC 241  
Db GAACAAGGACCAATTCGCTCACTAAGGTCTCTTTATCCAGATGTACCAATGTGGACC 3576  
QY 242 TAGACCTCGTAGGCTGGCGCGCTCCCAAGGTGCCGCTCATTAACACCATGCACTGCG 301  
Db TAGACCTCGTAGGCTGGCGCGCTCCCAAGGTGCCGCTCATTAACACCATGCACTGCG 3695  
QY 302 GCTCTCTCGGACCTTTTACCTGTGTACAGGCGACGCGCATGTCTATTCTGTGCCCGACGG 361  
Db GCTCTCTCGGACCTTTTACCTGTGTACAGGCGACGCGCATGTCTATTCTGTGCCCGACGG 3755  
QY 362 GTGATGGCAGGGGAGCCTGCTTTTCGCCCGCGGCTTATCTCTTACCTTGAAGGCTCTCTCG 421  
Db GTGATGGCAGGGGAGCCTGCTTTTCGCCCGCGGCTTATCTCTTACCTTGAAGGCTCTCTCG 3756  
QY 422 GAGGCGCTCTGCTGTGCCCGCGCAGACATCGGTAGGCATATTTCAGCGCGGTATGCA 481  
Db GAGGCGCTCTGCTGTGCCCGCGCAGACATCGGTAGGCATATTTCAGCGCGGTATGCA 3815  
QY 482 CCCGTGAGTGGCTAAGCGGTGAGCTTCACTCCCGTGTAGAGCTTATAGAGCAACCATGA 541  
Db CCCGTGAGTGGCTAAGCGGTGAGCTTCACTCCCGTGTAGAGCTTATAGAGCAACCATGA 3875  
QY 542 GGTCTCCCGGTGTTCTCAGACAACTCTCTCCCAACAGCAGTGCTCCAGAGCTTACCAAGTGG 601  
Db GGTCTCCCGGTGTTCTCAGACAACTCTCTCTCCCAACAGCAGTGCTCCAGAGCTTACCAAGTGG 3936  
QY 602 CCCACTGCATGCTCCCGCGCGGTGTAGAGCAACCAAGTCCCGCGCATACGCGAG 661  
Db CCCACTGCATGCTCCCGCGCGGTGTAGAGCAACCAAGTCCCGCGCATACGCGAG 4055  
QY 662 CTCAGGGCTCAAGGTGTGTGTCTCAACCCCTCGTTGTGCTGCAACAATGGGCTTTGGTG 721  
Db CTCAGGGCTCAAGGTGTGTGTCTCAACCCCTCGTTGTGCTGCAACAATGGGCTTTGGTG 4115  
QY 722 TTTACATGTCCAAGGCCCATGGGATGTATCTTAACATCAGAGCTGGGTGAGGACAATTA 781  
Db TTTACATGTCCAAGGCCCATGGGATGTATCTTAACATCAGAGCTGGGTGAGGACAATTA 4116  
QY 782 CTACTGGCAGCCCGATCAGTATTCACCTAGCGCAAGTTCCTTCCGACGCGGCTGTT 841  
Db CTACTGGCAGCCCGATCAGTATTCACCTAGCGCAAGTTCCTTCCGACGCGGCTGTT 4176  
QY 842 CAGGGGTGTCTTATGACATAATAATTTGTGAGAGTGGCACTTCCACGGATGCAATCCA 901  
Db CAGGGGTGTCTTATGACATAATAATTTGTGAGAGTGGCACTTCCACGGATGCAATCCA 4236  
QY 902 TCTTGGGATTTGGGCACTGTCTTGAACGAGAGACCGCGGGGCGAGACTGACTGTGC 961  
Db TCTTGGGATTTGGGCACTGTCTTGAACGAGAGACCGCGGGGCGAGACTGACTGTGC 4296  
QY 962 TCGCCACCGCTACCCCTCGGCTCCGTCACCTGTGCCCATCTTAACTCAGAGGAGTTG 1021  
Db TCGCCACCGCTACCCCTCGGCTCCGTCACCTGTGCCCATCTTAACTCAGAGGAGTTG 4415  
QY 1022 CTCTGTCCACTACCGGAGAGATCCCTTTTATGGCAAGGCTATTTCCCTTTGAAGCAATTA 1081



Db 4416 CTCTGTCCACACCGGAGAGATCCCTTTTACCGCAAGGCTATCCCTCGTGAAGTATCA 4475  
QY 1082 AGGGGGGAGACATCTCATCTTCTGCACTCAAGAGAGAGTGCAGAGCTCGCGCAA 1141  
Db 4476 AGGGGGGAGACATCTCATCTTCTGCACTCAAGAGAGAGTGCAGAGCTCGCGCAA 4535  
QY 1142 AATGTGTCGGTGGCGTCAATCGCGTCTTACTACCGCGCTTGATGTCCGTCA 1201  
Db 4536 AGCTGTGCGATTTGGGCATCAATCGCGTCTTACTACCGCGCTTGATGTCCGTCA 4595  
QY 1202 TCCGACCATGTTGAGCTGTGCTGTCGCACTGACCGCTCATGACCGCTTTACCG 1261  
Db 4596 TCCGACCATGTTGAGCTGTGCTGTCGCACTGACCGCTCATGACCGCTTTACCG 4655  
QY 1262 GCGACTTCGATTCGGTATAGACTGCAACACGCTGTGTACCCAGACAGTGCAGCC 1321  
Db 4656 GCGACTTCGATTCGGTATAGACTGCAACACGCTGTGTACCCAGACAGTGCAGCC 4715  
QY 1322 TTGACCTTACCTTACCATTTAGACAAATCAAGCTTCCCAAGATGCTGTCCCGTACTC 1381  
Db 4716 TTGACCTTACCTTACCATTTAGACAAATCAAGCTTCCCAAGATGCTGTCCCGTACTC 4775  
QY 1382 AAGCTCGGGTAGGACTGCGAGAGGAGCCAGGCATCTACAGATTTGTGGCACCGGGG 1441  
Db 4776 AAGCTCGGGTAGGACTGCGAGAGGAGCCAGGCATCTACAGATTTGTGGCACCGGGG 4835  
QY 1442 AGCGTCTCTGTCGATGTTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1501  
Db 4836 AGCGTCTCTGTCGATGTTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 4895  
QY 1502 CTTGGTATGAGCTTACGCGCGCGAGACACAGTGTAGGCTACGAGCATACATGAACACC 1561  
Db 4896 CTTGGTATGAGCTTACGCGCGCGAGACTACAGTGTAGGCTACGAGCATACATGAACACC 4955  
QY 1562 CGGACTTCCGCTGTGCAAGACATCTTGAATTTTGGAGGGCGCTTTTACGGTCTCA 1621  
Db 4956 CGGAGCTTCCGCTGTGCAAGACATCTTGAATTTTGGAGGGCGCTTTTACAGGCTCA 5015  
QY 1622 CCACATAGACGCCACTTCTTCCAGACAAAGAGAGTGGGGAACCTTCCCTATC 1681  
Db 5016 CTATATAGATGCCACTTCTTCCAGACAAAGAGAGTGGGGAACCTTCCCTATC 5075  
QY 1682 TGATAGCGTACCAAGCCACCGTGTGCGCTAGAGCTCAAGCCCTCCCGCTCGTGGGACC 1741  
Db 5076 TGATAGCGTACCAAGCCACCGTGTGCGCTAGAGCTCAAGCCCTCCCGCTCGTGGGACC 5135  
QY 1742 AGATGTGGAAGTGTGATCCGCTCTGAGCCCACTTCCATGGGCGCAACCTCTGCTAT 1801  
Db 5136 AGATGTGGAAGTGTGATCCGCTCTGAGCCCACTTCCATGGGCGCAACCTCTGCTAT 5195  
QY 1802 ATAGACTGGGCGTGTCCAGAAAGTCAACCTGACGCAACCCAGTCAACCAAGTATCA 1861  
Db 5196 ACAGACTGGGCGTGTTCAGAAAGTCAACCTGACGCAACCCAGTCAACCAAGTATCA 5255  
QY 1862 TGACATGTATGTGCGGTGACCTGAGAGTCTGACAGTACTCGGGTCTGTTGGCGGG 1921  
Db 5256 TGACATGTATGTGCGGTGACCTGAGAGTCTGACAGTACTCGGGTCTGTTGGCGGG 5315  
QY 1922 TTCTGCTGTTTGGCGCGTATTTGCTATTCACAGGCTCGGTGCTCATAGTAGGTAGGA 1981  
Db 5316 TCCTGCTGTTTGGCGCGTATTTGCTATTCACAGGCTCGGTGCTCATAGTAGGTAGGA 5375  
QY 1982 TTCTGCTGTTTGGCGCGTATTTGCTATTCACAGGCTCGGTGCTCATAGTAGGTAGGA 2041  
Db 5376 TCCTGCTGTTTGGCGCGTATTTGCTATTCACAGGCTCGGTGCTCATAGTAGGTAGGA 5435  
QY 2042 ATGAAATGGAAGTGTCT 2059  
Db 5436 ATGAAATGGAAGTGTCT 5453

RESULT 13

US-08-444-818-123/c  
; Sequence 123, Application US/08444818  
; Patent No. 6150087  
; GENERAL INFORMATION:  
; APPLICANT: Chien, David Y.  
; APPLICANT: Rucker, William J.  
; TITLE OF INVENTION: NABV Diagnostics and Vaccines  
; NUMBER OF SEQUENCES: 777  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Chiron Corporation  
; STREET: 4560 Horton Street  
; CITY: Emeryville  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 94608-2916  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent in Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/444,818  
; FILING DATE:  
; CLASSIFICATION: 424  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/08/403,590  
; FILING DATE: 14-MAR-1995  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Harbin, Alisa A.  
; REGISTRATION NUMBER: 33,895  
; REFERENCE/DOCKET NUMBER: 0110.002  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (508)359-3876  
; TELEFAX: (508)359-3885  
; INFORMATION FOR SEQ ID NO: 123:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 9185 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: cDNA  
; ANTI-SENSE: YES  
US-08-444-818-123

Query Match 86.6%; Score 1784.4; DB 3; Length 9185;  
Best Local Similarity 91.7%; Pred. No. 0;  
Matches 1887; Conservative 0; Mismatches 171; Indels 0; Gaps 0;  
QY 2 TGGCGCTATACAGGCTATGCCAGCAGCAAGGGGCTTTTGGGATGCAATCAACA 61  
Db 5790 TGGCGCTATACAGGCTATGCCAGCAGCAAGGGGCTTTTGGGATGCAATCAACA 5731  
QY 62 GCTTGNCCGCGGCAAAACCAGGTGAGGGTGCAGTTCAGATCGTCAACTGCTG 121  
Db 5730 GCTTAACTGCGCGGCAAAACCAGGTGAGGGTGCAGTTCAGATCGTCAACTGCTG 5671  
QY 122 CCAGAGCTTTCTTGGCAACTGCATTAAACGGGTGTTGGACTGTCTACATGGAGCCG 181  
Db 5670 CCAGAGCTTTCTTGGCAACTGCATTAAACGGGTGTTGGACTGTCTACATGGAGCCG 5611  
QY 182 GAACAGGACCAATTCGCTCACCTAAGGGTCTGTTATCCAGATGTACACCAATGTGAGC 241  
Db 5610 GAACAGGACCAATTCGCTCACCTAAGGGTCTGTTATCCAGATGTACACCAATGTGAGC 5551  
QY 242 AAGACTCTAGCTGGCGCTCCCAAGGTGCCGCTCATTAACACCATGCACTTGG 301  
Db 5550 AAGACTCTAGCTGGCGCTCCCAAGGTGCCGCTCATTAACACCATGCACTTGG 5491  
QY 302 GCTCTCTCGACCTTTTACCTGCTCAGGACGCGCGATGTCTTCTGTGCGCGGACGG 361  
Db 5490 GCTCTCTCGACCTTTTACCTGCTCAGGACGCGCGATGTCTTCTGTGCGCGGACGG 5431  
QY 362 GTGATGGCAGGGGACGCTGCTTTGCGCCCGGCTATCTTACTTTGAAAGGCTCCTCG 421



Db 5430 GTGATAGAGGGGACGCTGCTGTCGCCCGCGGCCATTTCTTACTTGAAGGCTCCTCGG 5371  
Qy 422 GAGGCCCTCTGCTGTGCCCGCAGGACATGCCGTAGGCATATTCAGAGCCGCGGTATGCA 481  
Db 5370 GGGGTCCGCTGTGTGTCGCCCGCGGGCACGCGTGGGCATATTTAGGGCCGCGGTGCA 5311  
Qy 482 CCGGTGAGTGGCTAAGGCGGTGGACTTTCATCCCGTAGAGCTTTAGAGACAACCATGA 541  
Db 5310 CCGGTGAGTGGCTAAGGCGGTGGACTTTCATCCCGTAGAGAACCTTAGAGACAACCATGA 5251  
Qy 542 GGTCCCGGCTGTCTCAGACAACCTCTCCCGCAGCAGTGGCCCGAGAGCTTACCAAGTGG 601  
Db 5250 GGTCCCGGCTGTCTCAGAGTAACTCTCTCCAGCAGTGGCCCGCAGAGCTTCCAGGTGG 5191  
Qy 602 CCACCTGATGCTCCCGCAGCGGTAAAGAGACCAAGGTCCCGCGCGCATACGAG 661  
Db 5190 CTACCTCCATGCTCCCGCAGCGGTAAAGAGACCAAGGTCCCGCGGTGATATGCGAG 5131  
Qy 662 CTACGGGCTACAGGTGCTGCTCAACCCCTCGTTCGCTGCAACAATGGGCTTTGGTG 721  
Db 5130 CTACGGGCTAAGGTGCTGAGTACTCAACCCCTCTGTTGCTGCAACACTGGGCTTTGGTG 5071  
Qy 722 CTACATGTCCAAAGGCCCATGGATTCCTAAACATCAGGACTGGGGTGAGGACAATTA 781  
Db 5070 CTACATGTCCAAAGGCTCATGGATCGATCCTAACATCAGGACCGGGGTGAGAACAAATTA 5011  
Qy 782 CTACTGGCAGCCGATCATGATTCACCTACCGGAAAGTTCCTGCGCAGCGGGGTGT 841  
Db 5010 CCACCTGGCAGCCCATCATGATTCACCTACCGGAAAGTTCCTGCGCAGCGGGGTGT 4951  
Qy 842 CAGGGGCTGTTATGACATAATTTGTGACAGTGGCCACTCCACGGATGCAACATCA 901  
Db 4950 CCGGGGCGGTTATGACATAATTTGTGACAGTGGCCACTCCACGGATGCAACATCA 4891  
Qy 902 TCTTGGGCAATGGCACTGCTTGGACCAAGCAGAGACCGCGGGGGGAGACTGACTGTGC 961  
Db 4890 TCTTGGGCAATGGCACTGCTTGGACCAAGCAGAGACTCGGGGGGAGACTGTTGTGC 4831  
Qy 962 TCGCAGCGGTACCCCTCGGGTCCGTCACGTGTCGCCCATCTTAACATCGAGGAGTTG 1021  
Db 4830 TCGCAGCGGTACCCCTCGGGTCCGTCACGTGTCGCCCATCTTAACATCGAGGAGTTG 4771  
Qy 1022 CTCTGTCCACTACGAGAGATCCCTTTTATGGCAAGCTATTTCCTTGAAGCAATTA 1081  
Db 4770 CTCTGTCCACTACGAGAGATCCCTTTTATGGCAAGCTATTTCCTTGAAGCAATTA 4711  
Qy 1082 AGGGGGGAGACATCTCATCTTCTGCCACTCAAAGAAAGTGCAGAGCTCGCGCAA 1141  
Db 4710 AGGGGGGAGACATCTCATCTTCTGCCACTCAAAGAAAGTGCAGAGCTCGCGCAA 4651  
Qy 1142 AACTGTGCGGTGGGGTCAATGCGGTGCTTACTACCGCGCTTGTATGTTCGCTCA 1201  
Db 4650 AGCTGTGCGATTTGGGCATCAATGCGGTGCTTACTACCGCGCTTGTATGTTCGCTCA 4591  
Qy 1202 TCCCGACAGTGGTGTGCTGCTGGCAACTGACGCGCTCATGACCGGCTTTACCG 1261  
Db 4590 TCCCGACAGTGGTGTGCTGCTGGCAACTGACGCGCTCATGACCGGCTTTACCG 4531  
Qy 1262 GCGACTTCGATTCGGTATAGATGCAACACGTGTGTACCCAGAGAGCTTCAGCC 1321  
Db 4530 GCGACTTCGATTCGGTATAGATGCAACACGTGTGTACCCAGAGAGCTTCAGCC 4471  
Qy 1322 TTGACCTTACCTTCAACATTTGACATCACTGCTTCCCGAGGATGCTGCTCCCGTAC 1381  
Db 4470 TTGACCTTACCTTCAACATTTGACATCACTGCTTCCCGAGGATGCTGCTCCCGTAC 4411  
Qy 1382 AACGTCCGGGTAGGACTGGCAGAGGGAAGCCAGGCATCTACAGATTTGTGGCACCGGGG 1441  
Db 4410 AACGTCCGGGTAGGACTGGCAGAGGGAAGCCAGGCATCTACAGATTTGTGGCACCGGGG 4351  
Qy 1442 AGGTCCTTCTGGCATGTTGACTGCTGCTCTCTGCGAGTGTATGACCGGGTGTG 1501

Db 4350 AGCGCCCTCGGCGCATGTTGACTGCTCGTCCGCTCTCTGTAGTGTATGACGCGAGTGTG 4291  
Qy 1502 CTTGGTATGAGCTTACGCCCGCGGAGACCAAGTACAGTACAGCATACATGACACCC 1561  
Db 4290 CTTGGTATGAGCTCACGCCCGCGGAGACTACAGTTAGGCTACGAGCGTACATGACACCC 4231  
Qy 1562 CGGACTTTCGGTGTGCCAAGACCATCTTGAATTTTGGAGGGGCTCTTTACGGGTCTCA 1621  
Db 4230 CGGGCTTTCGGTGTGCCAAGACCATCTTGAATTTTGGAGGGGCTCTTTACAGSCCTCA 4171  
Qy 1622 CCCATAGAGCGCCACTTCTATCCAGACAAGAGAGTGGGAAACCTTCCCTATC 1681  
Db 4170 CTCAATAGATGCCACTTCTATCCAGACAAGAGAGTGGGAGAACCTTCTTACC 4111  
Qy 1682 TGGTAGCTACCAAGCACCAGCTGCGCTAGAGCTCAAGCCCTCCCGCTCGTGGGACC 1741  
Db 4110 TGGTAGCTACCAAGCACCAGCTGCGCTAGAGCTCAAGCCCTCCCGCTCGTGGGACC 4051  
Qy 1742 AGATGTGAAGTGTGATTCGCTCTCAAGCCACCTCTCATGGGCCAACACCTCTGTAT 1801  
Db 4050 AGATGTGAAGTGTGATTCGCTCTCAAGCCACCTCTCATGGGCCAACACCTCTGTAT 3991  
Qy 1802 ATAGACTGGGCGCTGTCCAGAAATGAAGTCAAGTCAAGTCAAGTCAAGTCAAGTATCA 1861  
Db 3990 ACAGACTGGGCGCTGTTCAGAAATGAATCACCTGACGACCCAGTCAACAAATACATCA 3931  
Qy 1862 TGACATGATGTCGGCTGACCTGGAGTGTGTCAGAGTACCTGGTGTCTGTTGGCGGG 1921  
Db 3930 TGACATGATGTCGGCTGACCTGGAGTGTGTCAGAGTACCTGGTGTCTGTTGGCGGG 3871  
Qy 1922 TTTGGTGTCTTTGGCGCGCTATTGCTCTATCCAGAGTGTGCTGTATAGGTAGGA 1981  
Db 3870 TCTTGGTGTCTTTGGCGCGCTATTGCTCTATCCAGAGTGTGCTGTATAGGTAGGA 3811  
Qy 1982 TTGCTTGTCCGAAAGCCGCAATCATACCCGACAGGAGTCTCTTACCGGAGTTCG 2041  
Db 3810 TCGTCTTGTCCGGAAGCCGCAATCATACCTGACAGGGAAGTCTCTTACCGAGGTTTCG 3751  
Qy 2042 ATGAATGGAAGTGTCT 2059  
Db 3750 ATGATGGAAGTGTCT 3733

RESULT 14  
US-08-444-818-176  
; Sequence 176, Application US/08444818  
; Patent No. 6150087  
; GENERAL INFORMATION:  
; APPLICANT: Chien, David Y.  
; APPLICANT: Rutter, William J.  
; TITLE OF INVENTION: NANBV Diagnostics and Vaccines  
; NUMBER OF SEQUENCES: 777  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Chiron Corporation  
; STREET: 4560 Horton Street  
; CITY: Emeryville  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 94608-2916  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/444,818  
; FILING DATE:  
; CLASSIFICATION: 424  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/08/403,590  
; FILING DATE: 14-MAR-1995  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Harbin, Alisa A.



4536 AGTGTGCGCATTTGGGCATCAATCGCGTGGCTACTACCGCGTCTTGAAGTCCGTCA 4595  
 1202 TCCCGACAGTGGTGAAGTGTGCTGGTGGCAACTGAAGCCCTCATAGCCGGCTTTACGG 1261  
 4596 TCCCGACAGCGCGGATGTGCTGGTGGCAACCGATGCCCTCATAGCCGGCTTACCG 4655  
 1262 GCGACTTCGATTCCGGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1321  
 4656 GCGACTTCGACTCCGGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 4715  
 1322 TTGACCTTACCTTACCAATGAGACAATCAACGCTTCCCGAGGATGCTGTCTCCGTACTC 1381  
 4716 TTGACCTTACCTTACCAATGAGACAATCAACGCTTCCCGAGGATGCTGTCTCCGTACTC 4775  
 1382 AAGCTCGGGTAGGACTGGGAGGGAAGCCAGGATCTACAGATTTTGGGACCGGGGG 1441  
 4776 AAGCTCGGGTAGGACTGGGAGGGAAGCCAGGATCAACAGATTTTGGGACCGGGGG 4835  
 1442 AGGTCCTTCTGCGATGTTGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1501  
 4836 AGGTCCTTCTGCGATGTTGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 4895  
 1502 CTTGGTATGAGCTTACCGCGCGGAGACACACAGTTAGGCTACGAGCATATCAACACACC 1561  
 4896 CTTGGTATGAGCTTACCGCGCGGAGACACAGTTAGGCTACGAGCATATCAACACACC 4955  
 1562 CGGGAATTCCGGTGGCCAGACCATCTTGAATTTTGGGAGGGCGTCTTTAGGGTCTCA 1621  
 4956 CGGGAATTCCGGTGGCCAGACCATCTTGAATTTTGGGAGGGCGTCTTTACAGGCTCA 5015  
 1622 CCACATAGACGCCACTTCTTATCCAGACAAAGCAGAGTGGGAAACCTTCCCTATC 1681  
 5016 CTATATAGATGCCACTTCTTATCCAGACAAAGCAGAGTGGGAAACCTTCCCTATC 5075  
 1682 TGCTAGCTACCAAGCCACCGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1741  
 5076 TGCTAGCTACCAAGCCACCGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 5135  
 1742 AGATGGGAAGTCTGTATCCGCTCAAGCCACCTTCCATGGGCAACACCTTCTGCTAT 1801  
 5136 AGATGGGAAGTCTGTATCCGCTCAAGCCACCTTCCATGGGCAACACCTTCTGCTAT 5195  
 1802 ATAGATGGGCGCTGTCCAGAAATGAAGTCAACCTGACGCAACCGAGTCAACCAATATCA 1861  
 5196 ACAGATGGGCGCTGTTCAGAAATGAAGTCAACCTGACGCAACCGAGTCAACCAATATCA 5255  
 1862 TGACATGTATGTCCGCTGACCTGGAGTGTCTACGAGTACCTGGGTGCTGCTGGCGGG 1921  
 5256 TGACATGTATGTCCGCTGACCTGGAGTGTCTACGAGTACCTGGGTGCTGCTGGCGGG 5315  
 1922 TTCTGGTGTCTTGGCGCGTATTTGCTTATCCACAGGCTGCGTGTATAGTAGGTAGGA 1981  
 5316 TCCTGGTGTCTTGGCGCGTATTTGCTTATCCACAGGCTGCGTGTATAGTAGGTAGGA 5375  
 1982 TTGCTTGTCCGGAAGCCGGCAATCATACCCGACAGGGAAGTCTCTACCGGGAGTTCCG 2041  
 5376 TCGTCTTGTCCGGAAGCCGGCAATCATACCTGACAGGGAAGTCTCTACCGAGAGTTCCG 5435  
 2042 ATGAATGGAAGTGTCT 2059  
 5436 ATGAGATGGAAGTGTCT 5453

RESULT 15  
 US-09-881-239-2  
 ; Sequence 2, Application US/09881239  
 ; Patent No. 6630298  
 ; GENERAL INFORMATION:  
 ; APPLICANT: CHIEN, David Y.  
 ; APPLICANT: ARCANDEL, Phillip  
 ; APPLICANT: TANDESKE, Laura  
 ; APPLICANT: GEORGE-NASCIMENTO, Carlos  
 ; APPLICANT: COIT, Doris

; APPLICANT: MEDINA-SELBY, Angelica  
 ; TITLE OF INVENTION: HCV ANTIGEN/ANTIBODY COMBINATION ASSAY  
 ; FILE REFERENCE: 2302-16073 / PPI6073.003  
 ; CURRENT APPLICATION NUMBER: US/09/881,239  
 ; CURRENT FILING DATE: 2001-06-14  
 ; NUMBER OF SEQ ID NOS: 8  
 ; SOFTWARE: PatentIn Ver. 2.0  
 ; SEQ ID NO 2  
 ; LENGTH: 2058  
 ; TYPE: DNA  
 ; ORGANISM: Artificial Sequence  
 ; FEATURE:  
 ; OTHER INFORMATION: Description of Artificial Sequence:  
 ; OTHER INFORMATION: representative NS3/4a conformational antigen  
 ; NAME/KEY: CDS  
 ; LOCATION: (1)..(2058)  
 ; US-09-881-239-2

Query Match 86.4%; Score 1781.2; DB 4; Length 2058;  
 Best Local Similarity 91.6%; Pred. No. 0;  
 Matches 1885; Conservative 0; Mismatches 173; Indels 0; Gaps 0;  
 QY 1 ATGGCGCTATACAGGCTATGCCCAGAGACAAGGGGCTTTTGGGATGATTAATCACC 60  
 Db 1 ATGGCGCTATACAGGCTATGCCCAGAGACAAGGGGCTTTTGGGATGATTAATCACC 60  
 QY 61 AGCTTGACCGCGGACAAAACCAAGTGGAGGTGAGGTTTCAGATCGTCAACTGCT 120  
 Db 61 AGCTTGACCGCGGACAAAACCAAGTGGAGGTGAGGTTTCAGATCGTCAACTGCT 120  
 QY 121 GCCCAGAGCTTTCTTGGCAACCTGCAATTAACCGGGGTGTGTGGAGTGTCTACATGGAGCC 180  
 Db 121 GCCCAGAGCTTTCTTGGCAACCTGCAATTAACCGGGGTGTGTGGAGTGTCTACATGGAGCC 180  
 QY 181 GGAAACAGAGACATTTGCTGCTACCTTAAGGTCCTGTTATCCAGATGTACACCAATGTGAC 240  
 Db 181 GGAAACAGAGACATTTGCTGCTACCTTAAGGTCCTGTTATCCAGATGTATACCAATGTGAC 240  
 QY 241 CAAGACCTCTGAGGCTGCGCTCCCAAGGTGCGCTCAATTAACCAACCATGCACTTGC 300  
 Db 241 CAAGACCTCTGAGGCTGCGCTCCCAAGGTGCGCTCAATTAACCAACCATGCACTTGC 300  
 QY 301 GGCCTCTCGGACCTTTTACCTGCTACGAGGCAACCGGATGTCATTCCTGTGGCGCGGCG 360  
 Db 301 GGCCTCTCGGACCTTTTACCTGCTACGAGGCAACCGGATGTCATTCCTGTGGCGCGGCG 360  
 QY 361 GGTGATGACGAGGCGGAGCTGCTTTCGCGCGGCTATCTTACTTCAAGGCTCTCTCG 420  
 Db 361 GGTGATGACGAGGCGGAGCTGCTTTCGCGCGGCTATCTTACTTCAAGGCTCTCTCG 420  
 QY 421 GGAGGCTCTCTGCTGCTGCGCGGAGCATGCGGTAGGCATATTCAGAGCCGCGGTATGC 480  
 Db 421 GGAGGCTCTCTGCTGCTGCGCGGAGCATGCGGTAGGCATATTCAGAGCCGCGGTATGC 480  
 QY 481 ACCCGTGGAGTGGCTTAAGCGCGTGAATTCATTCCTGTAGAGAGCTTGAAGACCAATG 540  
 Db 481 ACCCGTGGAGTGGCTTAAGCGCGTGAATTCATTCCTGTAGAGAGCTTGAAGACCAATG 540  
 QY 541 AGGTCCCGGTGTTCTCAGACCACTCTCCACAGCAGTCCCGGAGGCTTACCAAGTG 600  
 Db 541 AGGTCCCGGTGTTCTCAGACCACTCTCCACAGCAGTCCCGGAGGCTTACCAAGTG 600  
 QY 601 GCCCAGCTGATGCTCCCAACCGGAGCGGTGAAGACCAACAGGTCCCGCGCATACGCA 660  
 Db 601 GCCCAGCTGATGCTCCCAACCGGAGCGGTGAAGACCAACAGGTCCCGCGCATATGCA 660  
 QY 661 GCTCAGGGCTACAAGGTGCTGCTCAACCCCTCGTGTGCTGCAACCAATGGGCTTTGGT 720  
 Db 661 GCTCAGGGCTACAAGGTGCTGCTCAACCCCTCGTGTGCTGCAACCAATGGGCTTTGGT 720  
 QY 721 GCTTACATGTCCCAAGGCGCATGATCTTAAACATCAGGACTGGGTGAGGCAAT 780  
 Db 721 GCTTACATGTCCCAAGGCGCATGATCTTAAACATCAGGACTGGGTGAGGCAAT 780

QY 781 ACTACTGGCAGCCCGATCAGTATTCCACCTACGGCAAGTTCCTTCCGACCGCGGGTGT 840  
Db |||||  
QY 781 ACCACTGGCAGCCCGATCAGTACTCACCTACGGCAAGTTCCTTCCGACCGCGGGTGT 840  
Db |||||  
QY 841 TCAGGGGGTCTTATGACATAAATAATTGTGACGAGTGCACCTCCACGGATGCAACATCC 900  
Db |||||  
QY 841 TCAGGGGGGGCTTATGACATAAATAATTGTGACGAGTGCACCTCCACGGATGCAACATCC 900  
Db |||||  
QY 901 ATCTTGGGCAATTCGGCACTGTCCTTTGACCAAGCAGAGACCGCGGGGCGAGACTGACTGTG 960  
Db |||||  
QY 901 ATCTTGGGCAATTCGGCACTGTCCTTTGACCAAGCAGAGACTGCGGGGGCGAGACTGCTGTG 960  
Db |||||  
QY 961 CTCGCCACCGCTACCCCTCCGGGCTCCGTCACTGTGCCCATCCTTAACATCGAGGAGGTT 1020  
Db |||||  
QY 961 CTCGCCACCGCCACCCCTCCGGGCTCCGTCACTGTGCCCATCCTTAACATCGAGGAGGTT 1020  
Db |||||  
QY 1021 GCTCTGTCCACTACCGGAGAGATCCCTTTTATGGCAAGGCTATTCCTTGGAGCAAT 1080  
Db |||||  
QY 1021 GCTCTGTCCACCGGAGAGATCCCTTTTACGGCAAGGCTATTCCTTGGAGTAATC 1080  
Db |||||  
QY 1081 AAGGGGGGAGACATCTCATCTTCTGCCACTCAAAGAAAGTGCGACGAGCTCGCGCGCA 1140  
Db |||||  
QY 1081 AAGGGGGGAGACATCTCATCTTCTGTCTTCAAAGAGAGTGCGAGACTCGCGCGCA 1140  
Db |||||  
QY 1141 AAACCTGGTCCGGTGGGGCGTCAATGCGGTGGCTTACTACCGGGGCCCTTGATGTGCCGT 1200  
Db |||||  
QY 1141 AAGCTGGTGGCATTTGGGCATCAATGCGGTGGCTTACTACCGGGGTCTTGACGGTGTCCGT 1200  
Db |||||  
QY 1201 ATCCGACGAGTGGTGAAGTGTGCTGTGGGAACTGAAGCCCTCATGACCGGCTTACC 1260  
Db |||||  
QY 1201 ATCCCGCCCATCGGCGATGTTGTGCTGTGGCAACCGATGCGCCCTCATGACCGGCTATACC 1260  
Db |||||  
QY 1261 GGGGACTTCGATTCGGGTGATAGACTGCAACAGCTGTGTCAACCCAGACAGTTCGACTTCAGC 1320  
Db |||||  
QY 1261 GGGGACTTCGATTCGGGTGATAGACTGCAATAGCTGTGTCAACCCAGACAGTTCGACTTCAGC 1320  
Db |||||  
QY 1321 CTTGACCCCTACCTTCAACCAATTGAGACAATCAGCTTTCCCGAGGATGCTGTCCCGTACT 1380  
Db |||||  
QY 1321 CTTGACCCCTACCTTCAACCAATTGAGACAATCAGCTTCCCGCAAGATGCTGTCTCCGCACT 1380  
Db |||||  
QY 1381 CAACGTCGGGGTAGGACTGCGCAGAGGAGCCAGGATCTACAGATTTGTGGACCCGGGG 1440  
Db |||||  
QY 1381 CAACGTCGGGGTAGGACTGCGCAGGGGAGCCAGGATCTACAGATTTGTGGACCCGGGG 1440  
Db |||||  
QY 1441 GAGCGTCTTCTGCGCATGTTTGAATGCTGTGCTCTGCGAGTGTATGACGGGGTGT 1500  
Db |||||  
QY 1441 GAGCGCCCTTCCGGCATGTTGCACTGCTCGCTCTCTGTGAGTGTATGACGAGGCTGT 1500  
Db |||||  
QY 1501 GCTTGGTATGAGCTTACGCCCGCGGAGACCAAGTTAGGCTACGAGCATACATGAACACC 1560  
Db |||||  
QY 1501 GCTTGGTATGAGCTACGCCCGCGGAGACTACAGTTAGGCTACGAGGTAGATGAACACC 1560  
Db |||||  
QY 1561 CCGGGACTTCCCGTGTGCGAAGACCATCTTGAATTTTGGAGGGGCTTTTACGGGTCTC 1620  
Db |||||  
QY 1561 CCGGGGCTTCCCGTGTGCGCAGGACCATCTTGAATTTTGGAGGGGCTTTTACAGGCTC 1620  
Db |||||  
QY 1621 ACCCAGATAGAGCCCACTTCTTATCCAGCAAAAGCAGAGTGGGAAACCTTTCCTAT 1680  
Db |||||  
QY 1621 ACTCATATAGATGCCCATTTCTTATCCAGCAAAAGCAGAGTGGGAGAACCTTTCCTTAC 1680  
Db |||||  
QY 1681 CTGGTAGCGTACCAAGCACCCTGTGCGCTAGAGCTCAAGCCCTCCCGCTCGTGGGAC 1740  
Db |||||  
QY 1681 CTGGTAGCGTACCAAGCACCCTGTGCGCTAGGCTCAAGCCCTCCCGCTCGTGGGAC 1740  
Db |||||  
QY 1741 CAGATGTGAAGTGTGATCCGTCTCAAGCCCAACCTTCATGGGCCAACACCTCTGCTA 1800  
Db |||||  
QY 1741 CAGATGTGAAGTGTGATTCGCTCAAGCCCAACCTTCATGGGCCAACACCCCTGCTA 1800  
Db |||||  
QY 1801 TATAGACTGGGCGCTGTCCAGATGAGTCACTGCGCAGCCAGTCAACCAAGTATATC 1860  
Db |||||  
QY 1801 TACAGACTGGGCGCTGTTCAGATGAATCACCCTGACGCCACCTGACCAAAATATATC 1860  
Db |||||

QY 1861 ATGACATGTATCTCGGCTGACCTGAGGTCGTCAAGAGTACCTTGGGTGCTCGTTGGCGGC 1920  
Db |||||  
QY 1861 ATGACATGTATCTCGGCTGACCTGAGGTCGTCAAGAGTACCTTGGGTGCTCGTTGGCGGC 1920  
Db |||||  
QY 1921 GTTCTGGCTGCTTGGCCGCGTATTGCCCTATCCACAGGCTGCGTGCATAGTAGGTAGG 1980  
Db |||||  
QY 1921 GTTCTGGCTGCTTGGCCGCGTATTGCCCTGTCAACAGGCTGCGTGCATAGTAGGTAGG 1980  
Db |||||  
QY 1981 ATTGTCTTGTCCGGAAAGCCGGCAATCATACCCGACAGGAAAGTCTCTACCGGGAGTTTC 2040  
Db |||||  
QY 1981 GTCTGTCTTGTCCGGGAGCCGGCAATCATACCTGACAGGAAAGTCTCTACCGGAGTTTC 2040  
Db |||||  
QY 2041 GATGAAATGGAAGAGTGC 2058  
Db |||||  
QY 2041 GATGAGATGGAAGAGTGC 2058  
Db |||||

Search completed: February 27, 2005, 22:43:12  
Job time : 376 secs

**THIS PAGE BLANK (USPTO)**

Result No.	Score	Query			DB	ID	Description
		Match	Length	Time			
C 1	46.8	2.3	509	9	CL252974	CL252974 ZMWB0060	
C 2	44.8	2.2	935	9	CNS006XK	AL056051 Drosophill	
C 3	44.4	2.2	834	4	BI956973	BI956973 HVSMen0000	
C 4	43.4	2.1	666	5	BM950570	BM950570 U1-M-EH0P	
C 5	42.8	2.1	1863	7	CV068974	CV068974 f2_new ch	
C 6	42.6	2.1	633	4	BI959933	BI959933 HVSMen002	
C 7	42.2	2.0	871	6	CD437613	CD437613 EL01N0503	
C 8	41.8	2.0	564	2	BE291962	BE291962 601085590	
C 9	41.8	2.0	672	4	BI155705	BI155705 6029004533	
C 10	41.8	2.0	935	5	BQ876236	BQ876236 AGENCOURT	
C 11	41.8	2.0	1020	6	BY117119	BY117119 BY117119	
C 12	41.8	2.0	1409	3	AK012576	AK012576 Mus muscu	
C 13	40.8	2.0	925	9	CNS0091P	AL053013 Drosophill	
C 14	40.4	2.0	533	4	BJ277484	BJ277484 BJ277484	
C 15	40.4	2.0	1749	9	CL948408	CL948408 OE1FS8005	
C 16	40.2	2.0	574	4	BJ208789	BJ208789 BJ208789	
C 17	40.2	2.0	629	4	BG906349	BG906349 TaLR1148G	
C 18	40.2	2.0	684	4	BI956133	BI956133 HVSMEm002	
C 19	40.1	1.9	697	7	CN944784	CN944784 011003AVB	
C 20	39.8	1.9	624	6	CD938118	CD938118 OV_109B20	
C 21	39.8	1.9	896	4	BQ899273	BQ899273 602836743	
C 22	39.6	1.9	424	5	BQ65971	BQ65971 HZ01L20u	
C 23	39.6	1.9	439	5	BQ65898	BQ65898 HZ01H12u	
C 24	39.6	1.9	420	5	BQ655954	BQ655954 HZ01L02u	





```

FEATURES
source
Location/Qualifiers
1. .666
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6"
/db_xref="taxon:10090"
/clone="IMAGE:5687271"
/tissue_type="whole brain"
/dev_stage="embryo 18.5 dpc"
/lab_host="DH10B (T1 phage resistant)"
/clone_lib="NIH_BMAP_EH0p"
/note="Organ: brain; Vector: pYX-Asc; Site 1: EcoR I;
Site 2: Not I; The library was constructed according to

```

ORIGIN

Query Match 2.1%; Score 42.8; DB 7; Length 1863;  
Best Local Similarity 48.4%; Pred. NO. 3.7;  
Matches 119; Conservative 0; Mismatches 127; Indels 0; Gaps 0;  
Qv 1126 GAGGAGCTCCGCCAAAACGTGTCGGGTGGGGCGTCAATGCGGTACTACCGCGC 1185

1523 GAGCCACAGGCGCATAGCGGACCCCTGCTCCATAGCGGCGCTCCCGCCATTGCC 1464  
 1186 CTTGATGTCGTCATCCGACCAAGTGTGAGCGTTGTCGTGCGCAACTGACGCGCTC 1245  
 1463 GCGGCGGTTACCGGCGCTCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTTC 1404  
 1246 ATGACCGCGCTTACCGCGCGCTTCGATTCGCTGATAGACTGCAACACGCTGTGTACCCAG 1305  
 1403 AAGCCACCTTCCCGCGCGCGCTGCGCGCGCTGCGCGCGCGCGCGCGCGCTTC 1344  
 1306 ACAGTCGATTCAGCTTACCGCTTACCTTACATGAGCAATACAGCTTCCCGAGAT 1365  
 1343 GTGCGCGCGCTCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGAT 1284  
 1366 GCTGTC 1371  
 1283 TCTGAC 1278

RESULT 6  
 BI959933  
 LOCUS  
 DEFINITION  
 HVSME002H07f Hordeum vulgare subsp. vulgare cDNA clone HVSME002H07f,  
 mRNA sequence.  
 ACCESSION  
 BI959933  
 VERSION  
 BI959933.1 GI:16311188  
 KEYWORDS  
 EST  
 SOURCE  
 Hordeum vulgare subsp. vulgare  
 ORGANISM  
 Hordeum vulgare subsp. vulgare  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
 Pooideae; Triticeae; Hordeum.  
 1 (bases 1 to 633)  
 Wing,R., Close,T.J., Kleinohs,A., Wise,R., Chin,A., Begum,D.,  
 Frisch,D., Atkins,M., Yu,Y., Henry,D., Palmer,M., Rambo,T.,  
 Simmons,J., Oates,R. and Main,D.  
 TITLE  
 Development of a genetically and physically anchored EST resource  
 for barley genomics: Morex rachis cDNA library  
 JOURNAL  
 Unpublished (2001)  
 COMMENT  
 Contact: Wing RA  
 Clemson University Genomics Institute  
 Clemson University  
 100 Jordan Hall, Clemson, SC 29634, USA  
 Tel: 864 656 7288  
 Fax: 864 656 4293  
 Email: rwing@clemson.edu  
 Total.hq bases = 513  
 Seq primer: AATTACCTCCTCAATAAGGG  
 High quality sequence stop: 595.  
 Location/Qualifiers  
 1. .633  
 /organism="Hordeum vulgare subsp. vulgare"  
 /mol\_type="mRNA"  
 /cultivar="Morex"  
 /sub\_species="vulgare"  
 /db\_xref="taxon:12509"  
 /clone="HVSME002H07f"  
 /tissue\_type="Rachis"  
 /lab\_host="TJG121"  
 /clone\_lib="Hordeum vulgare rachis EST library HVCDA0015  
 (normal)"]  
 /note="Vector: pBluescript SK(-); Site 1: EcoRI; Site 2:  
 XhoI; Plants were grown at Washington State University, and  
 Pullman, WA in a greenhouse, the rachises were excised, and  
 frozen in liquid nitrogen (Kleinohs lab). In the TJ Close  
 lab at the University of California, Riverside total RNA  
 was prepared, poly(A) was purified, one primary  
 unamplified cDNA library was made, and 1 million pfu were  
 in vivo excised to give pBluescript SK(-) cDNA phagemids  
 (Chin). Phagemids were plated and picked at the Clemson  
 University Genomics Institute (CUGI) (Begum, Palmer,

FEATURES  
 source  
 1. .633  
 /organism="Hordeum vulgare subsp. vulgare"  
 /mol\_type="mRNA"  
 /cultivar="Morex"  
 /sub\_species="vulgare"  
 /db\_xref="taxon:12509"  
 /clone="HVSME002H07f"  
 /tissue\_type="Rachis"  
 /lab\_host="TJG121"  
 /clone\_lib="Hordeum vulgare rachis EST library HVCDA0015  
 (normal)"]  
 /note="Vector: pBluescript SK(-); Site 1: EcoRI; Site 2:  
 XhoI; Plants were grown at Washington State University, and  
 Pullman, WA in a greenhouse, the rachises were excised, and  
 frozen in liquid nitrogen (Kleinohs lab). In the TJ Close  
 lab at the University of California, Riverside total RNA  
 was prepared, poly(A) was purified, one primary  
 unamplified cDNA library was made, and 1 million pfu were  
 in vivo excised to give pBluescript SK(-) cDNA phagemids  
 (Chin). Phagemids were plated and picked at the Clemson  
 University Genomics Institute (CUGI) (Begum, Palmer,

Frisch, Atkins and Wing). Plasmid DNA preparations, DNA  
 sequencing and sequence analysis were performed at CUGI  
 (Wing, Yu, Frisch, Henry, Simmons, Rambo, Main). The  
 sequence has been trimmed to remove vector sequence and  
 contains a minimum of 100 bases of phred value 20 or  
 above. For more details on library preparation and  
 sequence analysis see  
 http://www.genome.clemson.edu/projects/barley. To order  
 this clone see http://www.genome.clemson.edu/orders Also  
 see Close TJ, Wing R, Kleinohs A, Wise R (2001)  
 Genetically and physically anchored EST resources for  
 barley genomics. Barley Genetics Newsletter 31:29-30.  
 (http://wheat.pw.usda.gov/ggpages/bgn/31/cover.html)."

ORIGIN

Query Match 2.1%; Score 42.6; DB 4; Length 633;  
 Best Local Similarity 51.3%; Pred. No. 3.1; Indels 0; Gaps 0;  
 Matches 99; Conservative 0; Mismatches 94; Gaps 0;  
 QY 1096 CTCATCTTCTGCCACTCAAGAGAGAGTGCAGAGCTCCGCCAAATCTGTCGGGTG 1155  
 Db 405 CTCTTCAACTGCTCCAGCGGACAGAACTGCGCGGCGCCCAACGGGACGCGGTTCCGGGCG 464  
 QY 1156 GCGCTCAATGCGGTGTTACTACCGGCGCTTGTGTCCTCATCCCGGACAGTGTG 1215  
 Db 465 AGCATCAACAACGCTCTTCTGCTCCCTCCACCGTCTCCATCTTCCAGGACACTAC 524  
 QY 1216 GACGTTGTCGTGCGGCAACTGACGCGCTCATACCGGCTTACCGGCGACTTCGATTCG 1275  
 Db 525 CAGGCGATGCG 584  
 QY 1276 GTGATAGACTGCA 1288  
 Db 595 CAGTTCGACTACA 597

RESULT 7

CD437613  
 LOCUS  
 DEFINITION  
 EL01N0503A05.b Endosperm\_5 Zea mays cDNA, mRNA sequence.  
 ACCESSION  
 CD437613  
 VERSION  
 CD437613.1 GI:31353256  
 KEYWORDS  
 EST.  
 SOURCE  
 Zea mays  
 ORGANISM  
 Zea mays  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD  
 clade; Panicoideae; Andropogoneae; Zea.  
 1 (bases 1 to 871)  
 Lai,J., Dey,N., Kim,C.S., Bharti,A.K., Rudd,S., Mayer,K.F.,  
 Larkins,B., Becraft,P. and Messing,J.  
 TITLE  
 Characterization of the maize endosperm transcriptome and its  
 comparison to the rice genome  
 JOURNAL  
 Genome Res. 14 (10), 1932-1937 (2004)  
 COMMENT  
 Contact: Lai, Jinsheng  
 Dr. Joachim Messing's lab  
 Wakeman Institute, Rutgers University  
 190 Frelinghuysen Rd., Piscataway, NJ 08854, USA  
 Tel: 732-445-3801  
 Fax: 732-445-5735  
 Email: jlai@wakeman.rutgers.edu  
 Seq primer: T3  
 Location/Qualifiers  
 1. .871  
 /organism="Zea mays"  
 /mol\_type="mRNA"  
 /cultivar="W22"  
 /db\_xref="taxon:4577"  
 /tissue\_type="Endosperm of 7-23DAP"  
 /clone\_lib="Endosperm\_5"  
 /note="Vector: pBluescript SK-; Site 1: EcoRI; Site 2:  
 XhoI"

FEATURES

source

ORIGIN

```
Query Match      2.0%; Score 42.2; DB 6; Length 871;
Best Local Similarity 48.9%; Pred. No. 4.4;
Matches 113; Conservative 0; Mismatches 118; Indels 0; Gaps 0;

QY 1144 CTGGTCGGTGGCGCTCAATGCGTGGCTTACTACCGCGCCCTTGTATGTGTCGGTCATC 1203
Db 151 CTGCGCGCGAGTGGCTCGCGGACATTTGCGAGTACCTCGCGGACCGCTTTCTCTCGTCGAC 210
QY 1204 CCGACCACTGCTGACGTTGTGCTGCGTGGCACTGAGCCCTCATGACCGGCTTTACCGGC 1263
Db 211 CCGCGCTTCTCGCGACATCGAGTCACTGCGAGGCTCAAGGACCCCTACCGGC 270
QY 1264 GACTTCGATTGCTGATGAGTCAACACAGTGTGTACCCAGAGTCGACTTCAGCCTT 1323
Db 271 GACTTCGACGCGCGGCTCTCGCGCGCTCGCTCGTATCTCGAGGCGGAGTTCCGCGC 330
QY 1324 GACCTACCTTACCATTTGAGACATCACGCTTCCCGAGATGCTCTCC 1374
Db 331 CTCTCGCGCACTCCGCGCGCTCGCGATGCGGAGCCGCGCGCGC 381

RESULT 8
BE291962/c
LOCUS      601085590F1 NCI_CGAP_Mam6 Mus musculus cDNA clone IMAGE:3499726 5',
DEFINITION mRNA sequence.
ACCESSION BE291962
VERSION    BE291962.1 GI:9174040
KEYWORDS   EST.
SOURCE     Mus musculus (house mouse)
ORGANISM   Mus musculus

REFERENCE   1 (bases 1 to 564)
AUTHORS    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
TITLE       Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
JOURNAL    NIH-MGC http://mgc.nci.nih.gov/.
COMMENT     National Institutes of Health, Mammalian Gene Collection (MGC)
            Unpublished (1999)
            Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Jeffrey Green M.D.
            CDNA Library Preparation: Life Technologies, Inc.
            DNA Sequencing by: Incyte Genomics, Inc.
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM8557 row: c column: 23
            High quality sequence stop: 561.
            Location/Qualifiers
                1..564
                /organism="Mus musculus"
                /mol_type="mRNA"
                /strain="FVB/N"
                /db_xref="taxon:10090"
                /clone="IMAGE:3499726"
                /sex="female, virgin"
                /tissue_type="infiltrating ductal carcinoma"
                /dev_stage="5 months"
                /lab_host="DH10B"
                /clone_lib="NCI_CGAP_Mam6"
                /notes="Organ: mammary; Vector: pCMV-SPORT6; Site 1: SalI;
                Site 2: NotI; Cloned unidirectionally. Primer: Oligo dt.
                Library constructed by Life Technologies. Investigator
                providing samples: Jeffrey Green, M.D., NIH"

ORIGIN
Query Match      2.0%; Score 41.8; DB 2; Length 564;
Best Local Similarity 53.3%; Pred. No. 5;
Matches 88; Conservative 0; Mismatches 77; Indels 0; Gaps 0;

QY 687 CAACCCCTCGTTGCTGCAACAATGGGCTTTGGTGTATGTTCAAGGCCCATGGGAT 746
Db 397 CAAGGCCACCTGTCTGGAGCCATGGGCTCTCGAGTGTCTCGCTCAGAGCCACAGGGCT 338
QY 747 TGATCTTAACATCAGGACTGGGGTGAGGACAATTAATCTACTGGAGCCCATCAGTATTC 806
Db 337 GGACACAGAGGGCAGCTCTCTCGGGTAGGCAAGGTCACAGGAGCCTGGGGCTAGTGCC 278
QY 807 CACCTACGCGCAAGTTCTTTGCCGAGCGGGGTTTCAGGGGTGC 851
Db 277 CGGCTCTGGAGCAGCTGTCAGCGCGCGGTGGAGTTTCAGCTGAGGC 233

Query Match      2.0%; Score 41.8; DB 4; Length 673;
Best Local Similarity 53.3%; Pred. No. 5.3;
Matches 88; Conservative 0; Mismatches 77; Indels 0; Gaps 0;

QY 687 CAACCCCTCGTTGCTGCAACAATGGGCTTTGGTGTATGTTCAAGGCCCATGGGAT 746
Db 397 CAAGGCCACCTGTCTGGAGCCATGGGCTCTCGAGTGTCTCGCTCAGAGCCACAGGGCT 338
QY 747 TGATCTTAACATCAGGACTGGGGTGAGGACAATTAATCTACTGGAGCCCATCAGTATTC 806
Db 337 GGACACAGAGGGCAGCTCTCTCGGGTAGGCAAGGTCACAGGAGCCTGGGGCTAGTGCC 278
QY 807 CACCTACGCGCAAGTTCTTTGCCGAGCGGGGTTTCAGGGGTGC 851
Db 277 CGGCTCTGGAGCAGCTGTCAGCGCGCGGTGGAGTTTCAGCTGAGGC 233
```



```
/tissue_type="whole body"
/dev_stage="11 days embryo"
/clone_lib="RIKEN full-length enriched, 11 days embryo
whole body"

ORIGIN
Query Match 2.0%; Score 41.8; DB 6; Length 1020;
Best Local Similarity 53.3%; Pred. No. 5.9;
Matches 88; Conservative 0; Mismatches 77; Indels 0; Gaps 0;

QY 687 CAACCCCTCGTGTGCAACATGAGTGGCTTTGGTCTTACATGTCACAGGCCCATGGAT 746
Db 352 CAAGGCCACTGTGCTGGAGCCATGGGCTCTGAGTGTCTCTCCTCAGAGCCACAGGGCT 293

QY 747 TGATCTTAACATCAGCAGTGGGTGAGGACAAATTACTTCTGCGAGCCCGATCACGTATTC 806
Db 292 GGACACAGGGCAGCTCTCGGTAGGCAAGTCCACAGGAGCGCTGGGGCTAGTGCC 233

QY 807 CACCTACGGCAAGTCTTGTGCGACCGCGGTGTTTCAGGGGGTGC 851
Db 232 CGCTCTGGAGCAGCTGCACGCCCGCGTGGAGTTCAGCTGAGGC 188

RESULT 12
AK012576/c
LOCUS
DEFINITION
AK012576 1409 bp mRNA linear HTC 03-APR-2004
Mus musculus 11 days embryo whole body cDNA, RIKEN full-length
enriched library, clone:2700087H15 product:hypothetical protein,
full insert sequence.
ACCESSION
AK012576
VERSION
AK012576.1 GI:12849408
KEYWORDS
HTC; CAP trapper.
SOURCE
Mus musculus (house mouse)
ORGANISM
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
1 Carninci, P. and Hayashizaki, Y.
High-efficiency full-length cDNA cloning
Meth. Enzymol. 303, 19-44 (1999)
JOURNAL
99279253
MEDLINE
10349636
PUBMED
2 Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes
Genome Res. 10 (10), 1617-1630 (2000)
JOURNAL
20499374
MEDLINE
11042159
PUBMED
3 Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,
Konno, H., Akiyama, J., Nishi, K., Katsunai, T., Tashiro, H., Itoh, M.,
Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,
Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,
Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohata, E., Watanabe, M.,
Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J.,
Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multicapillary sequencer
Genome Res. 10 (11), 1757-1771 (2000)
JOURNAL
20530913
MEDLINE
11076861
PUBMED
4 The RIKEN Genome Exploration Research Group Phase II Team and the
FANTOM Consortium.
Functional annotation of a full-length mouse cDNA collection
Nature 409, 685-690 (2001)
JOURNAL
5 The FANTOM Consortium and the RIKEN Genome Exploration Research
Group Phase I & II Team.
Analysis of the mouse transcriptome based on functional annotation
of 60,770 full-length cDNAs
```

```

747 TGATCTTAACATCAGGACTGGGGTGAGGACAATTACTACTCGCAGCCCGATCACGATTTC 806
Ddb
292 GGACACAAAGGGGCAGTCTCTCGGGTAGGCCAAGGTCCACAGCGACCTGGGGCTAGTGCC 233
Qy
807 CACTACGGCAAGTTCTTTCGCCAGCGGGGTGTTTCAGGGGTGC 851
Ddb
232 CGGCTCTGGAGCAGCTGCGAGCCCGCGTGGGAGTTTCAGCTGAGGC 188

RESULT 13
CONS0091P/C
LOCUS
DEFINITION
  CNS0091P          925 bp      DNA           linear       GSS 03-JUN-1999
  Drosophila melanogaster genome survey sequence TET3 end of BAC #
  BACR19D16 of RPCI-98 library from Drosophila melanogaster (fruit
  fly), genomic survey sequence.
  AL053013
  AL053013.1 GI:4934461
  GSS.
  Drosophila melanogaster (fruit fly)
  Drosophila melanogaster
  Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
  Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
  Ephydroidea; Drosophilidae; Drosophila.
  1 (bases 1 to 925)
  Genoscope.
Direct Submission
Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
Determination of this BAC-end sequence was carried out as part of a
collaboration with the Berkeley Drosophila Genome Project (BDGP).
The BDGP is constructing a physical map of the Drosophila
melanogaster genome using these BACs. For further information
please see http://www.fruitfly.org/The BDGP Drosophila
melanogaster BAC library was prepared by Kazutoyo Osoegawa and
Aaron Mammosser in Pietter de Jong's laboratory in the Department of
Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,
NY. The library is named RPCI-98 and was constructed by partial
EcoRI digestion of Drosophila DNA provided by the BDGP from the
isogenic strain y2; cn bw sp, the same strain used for the BDGP's
PI and EST libraries. A more detailed description of the library
and how to hybridize individual BAC clones, the entire library, or
filers for hybridization from the BACPAC Resource Center can be
found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.
Location/Qualifiers
    1..925
     /organism="Drosophila melanogaster"
     /mol_type="genomic DNA"
     /db_xref="taxon:7227"
     /clone="BACR19D16"
     /clone_lib="RPCI-98"
     /note="end : TET3"

FEATURES
             source
Query Match          2.0%; Score 40.8; DB 9; Length 925;
Best Local Similarity 12.9%; Pred.No.11;
Matches 47; Conservative 159; Mismatches 158; Indels 0; Gaps 0

Qy   346  CTGTGTCGCCGACGGGTGATGGCAGGGCAGCGCTGCTTCGCCCGGCTATCTCTTAC 405
Db   916  SCSSBSCSSSMSTSSNSSBCSSBSBSSTSSMSSBSBSBSGSSSSSGTSSACV 857
Qy   406  TTCAAAGGCTCTTCGGGAGGGCCCTCTGCTGTGCCCGCCAGGACATCCGCTAGCATATTC 465
Db   856  KCVASSCCCGCGMABCMWCSSSSCGSASARGVKVASGGAGKRGGGGCGASASHS 797
Qy   466  AGAGCCCGGTATGACCCTGTGAGTGCTGAAGCGGTGGACTTCATCCCCTAGAGAC 525
Db   796  SSACBSSSSSCSASWSASSSSSSRSRRSGGGAGAGSSASSRSSSSSSASAGSVSS 737
Qy   526  TTAGACACAAACATAGGTCCCGCGTGTCTCAGACAACTCTCCCCACACAGCTGCC 585

```

```

VERSION      CL948408.1  GI:52360417
KEYWORDS     GSS.
SOURCE       Oryza sativa (indica cultivar-group)
ORGANISM     Oryza sativa (indica cultivar-group)
              Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
              Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
              Ehrhartoideae; Oryzeae; Oryza.
REFERENCE    1 (bases 1 to 1749)
AUTHORS      Ma, L., Wang, J., Chen, C., Liu, X., Su, N., Li, L., Wang, X., Cao, M.,
              Jiao, Y., Sun, N., Zhang, X., Bao, J., Sun, D., Zhao, H., Yuan, L.,
              Wong, G.K.S., Deng, X.W. and Wang, J.
TITLE        An analysis of transcriptional regulation of the rice genome and
              its comparison to Arabidopsis
JOURNAL      Unpublished (2004)
COMMENT      Department of Bioinformatic
              Beijing Institute of Genomics
              Chinese Academy of Sciences, Beijing 101300, China
              Tel: 86-10-80481559
              Fax: 86-10-80488676
              Email: chenchen@genomics.org.cn
              Rice genomic sequence.
              Class: exon-trapped.
FEATURES     source
              1..1749
               /organism="Oryza sativa (indica cultivar-group)"
               /mol_type="genomic DNA"
               /db_xref="taxon:39946"
               /clone_lib="Oryza sativa Express Library"
               /note="Oryza sativa exon trapped genomic sequences"

ORIGIN
Query Match      2.0%; Score 40.4; DB 9; Length 1749;
Best Local Similarity 63.3%; Pred. No. 17;
Matches 62; Conservative 0; Mismatches 36; Indels 0; Gaps 0;

QY 1325 ACCCTACCTTCACCATTTAGACAAATCACGCTTCCCGAGGATGCTGTCTCCGCTACTCAAC 1384
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 539 ACTCTCCCGCACCGGTGGCGCGCGGCTACCTACCCCGAGTCTGCCCGCGCTCTAAGC 598
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 1385 GTCCGGGTAGGACTGCGAGAGGAGCCAGGCATCTAC 1422
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 599 GACGAGATAGGACAGGCAAGGAACACGCGCGCTCC 636
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Search completed: February 27, 2005, 22:36:51
Job time : 6109 secs

```



**THIS PAGE BLANK (USPTO)**